

Aus der Klinik für Kardiologie und Pneumologie
(Prof. Dr. med. G. Hasenfuß)
der Medizinischen Fakultät der Universität Göttingen

**Effektivität von Kleingruppenunterricht
im Vergleich zu einer
Computersimulation in einem ‚*Blended
learning*‘- Szenario im Medizinstudium**

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von

Angéline-Charline Middeke

aus

Höxter

Göttingen 2018

Dekan:	Prof. Dr. rer. nat. H. K. Kroemer
Referent/in	Prof. Dr. T. Raupach
Ko-Referent/in:	PD Dr. S. Sennhenn-Kirchner

Datum der mündlichen Prüfung: 30.04.2019

Hiermit erkläre ich, die Dissertation mit dem Titel „Effektivität von Kleingruppenunterricht im Vergleich zu einer Computersimulation in einem ‚Blended learning‘- Szenario im Medizinstudium“ eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den
.....
(Unterschrift)

Die Ergebnisse dieser Arbeit wurden teilweise veröffentlicht:

Middeke A, Schuelper N, Raupach T: Introducing the virtual A&E department EMERGE to undergraduate medical education: a prospective trial. Round table #10R3 auf der AMEE Konferenz 2017, Helsinki.

Middeke A, Anders S, Schuelper M, Raupach T, Schuelper N (2018): Training of clinical reasoning with a Serious Game versus small-group problem-based learning: A prospective study. PLOS ONE 13, e0203851.

Inhaltsverzeichnis

Inhaltsverzeichnis	1
Abbildungsverzeichnis	4
Tabellenverzeichnis	4
Abkürzungsverzeichnis.....	5
1 Einleitung	6
1.1 Der Aufbau des Medizinstudiums an der Universitätsmedizin Göttingen.....	6
1.2 Prinzipien der Curricularentwicklung.....	7
1.2.1 Zielgerichtete Curricularplanung	7
1.2.2 ‚Outcome-based education‘	8
1.2.3 Göttinger Lernzielkatalog	9
1.3 ‚Clinical reasoning‘	10
1.3.1 Zwei-Prozess-Theorien	10
1.3.2 Der diagnostische Entscheidungsprozess	11
1.3.3 Zentrale Elemente des ‚clinical reasoning‘	12
1.3.4 Konsequenzen für die medizinische Lehre	12
1.4 Lehrformen	13
1.4.1 Die Vorlesung.....	14
1.4.2 Das Seminar	14
1.4.3 Das Problemorientierte Lernen	15
1.4.4 Der Unterricht am Krankenbett	16
1.4.5 Simulationen.....	18
1.4.6 ‚Serious Games‘	18
1.5 Prüfungen	23
1.5.1 Prüfungskonsequenzen.....	23
1.5.2 Prüfungsformate.....	24
1.6 Ziele dieser Arbeit	25
2 Methoden	27
2.1 Studiendesign	27
2.2 Organisation und Durchführung.....	27
2.2.1 Probandenrekrutierung.....	28
2.3 Lehrformen	29
2.3.1 Kleingruppenunterricht POLEMA	29
2.3.2 Computersimulation EMERGE.....	30
2.4 Die formative Abschlussprüfung.....	36
2.4.1 Organisation der Abschlussprüfung.....	36
2.4.2 Fragebögen	37

2.4.3	„Key Feature“-Prüfung.....	38
2.4.4	EMERGE-Prüfung.....	39
2.5	Fokusgruppengespräche.....	41
2.5.1	Vorbereitung und Durchführung	41
2.5.2	Transkription und Auswertung.....	42
2.6	Statistische Analyse	42
3	Ergebnisse.....	44
3.1	Charakterisierung der Studiengruppe.....	44
3.2	Anwesenheit bei den Seminaren	45
3.3	Studienfrage 1: Differentialdiagnostische/-therapeutische Kompetenzen	45
3.4	Studienfrage 2: Format- und/oder Fallspezifität der erlernten Kompetenzen	47
3.4.1	Formatspezifität.....	47
3.4.2	Fallspezifität	49
3.5	Studienfrage 3: Prozedurales Vorgehen.....	52
3.6	Studienfragen 4 und 5: Ergebnisse der Fragebögen und der Fokusgruppengespräche.....	55
3.6.1	Ergebnisse der Fragebögen.....	55
3.6.2	Ergebnisse der Fokusgruppengespräche	58
4	Diskussion.....	64
4.1	Die wesentlichen Ergebnisse.....	64
4.2	Effektivität von EMERGE hinsichtlich differentialdiagnostischer und – therapeutischer Kompetenzen (Studienfrage 1)	65
4.2.1	Vergleich mit der bestehenden Literatur	67
4.2.2	Einsatz von „Key Feature“-Fällen zur Prüfung differentialdiagnostischer und – therapeutischer Kompetenzen	70
4.2.3	„Blended learning“	72
4.3	Format- und/oder Fallspezifität der erlernten Kompetenzen (Studienfrage 2).....	73
4.3.1	Formatspezifität.....	73
4.3.2	Fallspezifität	74
4.4	Prozedurales Vorgehen (Studienfrage 3).....	75
4.4.1	Priorisierung von Labor/apparativen Untersuchungen im Fall „NSTEMI“	78
4.5	Stressempfinden bei der Behandlung mehrerer virtueller Patienten (Studienfrage 4)	79
4.6	EMERGE als geeignetes Lehrformat für den Erwerb ärztlicher Kompetenzen (Studienfrage 5).....	81
4.6.1	Fokusgruppengespräche zur Evaluation von Meinungen bezüglich des Einsatzes von EMERGE in der Lehre.....	83
4.7	Studiendesign	84
4.8	Ausblick.....	86
4.8.1	Effektiver Einsatz von „Serious Games“ in der Lehre.....	86
4.8.2	Langzeitretention erworbener Kompetenzen.....	87
4.8.3	Transfer von Wissen und Fähigkeiten aus der Virtualität in die Realität	88
5	Zusammenfassung.....	89

6	Anhang	91
7	Literaturverzeichnis	99

Abbildungsverzeichnis

Abbildung 1: Bildschirmaufnahmen von EMERGE	32
Abbildung 2: Zusammensetzung der Studiengruppe.....	44
Abbildung 3: Ergebnisse der ‚Key Feature‘-Abschlussprüfung (Vergleich EMERGE und POLEMA)	46
Abbildung 4: Ergebnisse der ‚Key Feature‘-Abschlussprüfung hinsichtlich der in den Lehrformen zuvor thematisierten Inhalte (Vergleich EMERGE und POLEMA).....	46
Abbildung 5: Ergebnisse der ‚Key Feature‘-Abschlussprüfung im Hinblick auf die inhaltlichen Schwerpunkte der drei EMERGE-Gruppen (Vergleich EMERGE A, B und AB).....	51

Tabellenverzeichnis

Tabelle 1: Modifiziert entnommen aus dem Göttinger Lernzielkatalog, Teil 2: Gesundheitsstörungen	9
Tabelle 2: Inhalte von POLEMA im Sommersemester 2016	30
Tabelle 3: EMERGE-Fälle, die im Rahmen der Studie eingesetzt wurden.	33
Tabelle 4: Einteilung der EMERGE-Fälle in acht Fallgruppen.....	34
Tabelle 5: Verteilung der EMERGE-Fallgruppen auf die einzelnen Termine der drei EMERGE-Gruppen.....	35
Tabelle 6: Exposition der Gruppen gegenüber den Inhalten der ‚Key Feature‘-Fälle.....	38
Tabelle 7: Festgelegte Variablen zur Auswertung der EMERGE-Fälle	40
Tabelle 8: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich der Globalbewertung und der inhaltlichen Fähigkeiten (Vergleich POLEMA und EMERGE)	47
Tabelle 9: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich prozeduraler Fähigkeiten (Vergleich POLEMA und EMERGE).....	48
Tabelle 10: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich der inhaltlichen Fähigkeiten (Vergleich EMERGE-Gruppen A, B und AB).....	49
Tabelle 11: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich prozeduraler Fähigkeiten (Vergleich EMERGE-Gruppen A, B und AB).....	49
Tabelle 12: Ergebnisse der ‚Key Feature‘-Abschlussprüfung (Vergleich EMERGE-Gruppen A, B und AB)	51
Tabelle 13: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich des prozeduralen Vorgehens (Vergleich POLEMA und EMERGE)	54
Tabelle 14: Ergebnisse der Fragebögen hinsichtlich soziodemographischer Daten, Vorerfahrung im Rettungswesen, Nutzung digitaler Medien und Stressempfinden (Vergleich POLEMA und EMERGE).....	55
Tabelle 15: Ergebnisse der Fragebögen hinsichtlich der Motivation zur Teilnahme an EMERGE und der inhaltlichen Bewertung der virtuellen Notaufnahme (nur EMERGE).....	57
Tabelle 16: Ergebnisse der Fokusgruppengespräche – EMERGE im Vergleich zu anderen Lehrformen.....	60
Tabelle 17: Ergebnisse der Fokusgruppengespräche – Inhaltliche Bewertung der virtuellen Notaufnahme.....	61
Tabelle 18: Ergebnisse der Fokusgruppengespräche – Organisation des Lehrformates	62
Tabelle 19: Ergebnisse der Fokusgruppengespräche – Erlebte Emotionen während der EMERGE-Termine.....	63
Tabelle 20: Ergebnisse der Fokusgruppengespräche – Technische Aspekte.....	63

Abkürzungsverzeichnis

3D	Dreidimensional
ÄApprO	Approbationsordnung für Ärzte
DiPS	Digitales Prüfungs- und Schulungszentrum
,E-Learning‘	, <i>Electronic Learning</i> ‘
FG	Fokusgruppengespräch
GMA	Gesellschaft für medizinische Ausbildung
,OSCE‘	, <i>Objective Structured Clinical Examination</i> ‘
PJ	Praktisches Jahr
POLEMA	Problem orientiertes Lernen – elektronische Medien assistiert
STÄPS	Studentisches Trainingszentrum Ärztlicher Praxis und Simulation
TXT-Datei	Textdatei
UMG	Universitätsmedizin Göttingen

1 Einleitung

1.1 Der Aufbau des Medizinstudiums an der Universitätsmedizin Göttingen

Die ärztliche Ausbildung umfasst ein Studium der Humanmedizin mit einer Regelstudienzeit von sechs Jahren und drei Monaten (Güntert 2003). In einem Regelstudiengang, wie er an der Universitätsmedizin Göttingen (UMG) angeboten wird, ist das Studium in zwei Abschnitte unterteilt. Der erste Studienabschnitt umfasst vier Fachsemester mit naturwissenschaftlichem Schwerpunkt. Diese so genannte Vorklinik wird nach zwei Jahren mit dem Ersten Abschnitt der Ärztlichen Prüfung („M1/Physikum“) abgeschlossen. Zur Prüfungsanmeldung muss darüber hinaus eine Ausbildung in erster Hilfe sowie ein dreimonatiger Krankenpflagedienst nachgewiesen werden. Im zweiten, klinischen Studienabschnitt, welcher sechs Fachsemester mit wissenschaftlich-praktischem Schwerpunkt umfasst, müssen Leistungsnachweise in insgesamt 21 Hauptfächern, 14 Querschnittsbereichen, einem Wahlfach und fünf Blockpraktika erbracht werden.

An der UMG ist die Lehre im klinischen Studienabschnitt in 21 interdisziplinären Modulen organisiert. Im Rahmen eines „themen- und organbezogenen Blockunterrichts“ (Universitätsmedizin Göttingen – Klinischer Studienabschnitt Humanmedizin) werden Lehrinhalte fächerübergreifend vermittelt, um ein leichteres Verständnis zu gewährleisten und eine interdisziplinäre Herangehensweise an klinische Fragestellungen zu fördern. Im ersten und zweiten klinischen Semester werden auf diese Weise die Grundlagen von Krankheitslehre und Diagnostik gelehrt, um auf die Vermittlung spezifischer Krankheitslehre in den klinischen Semestern drei bis fünf vorzubereiten. Das sechste klinische Semester dient zur Vertiefung des zuvor erworbenen Wissens (Universitätsmedizin Göttingen – Curriculum Klinik). Im Rahmen des im Sommersemester 2016 neu eingeführten und im sechsten klinischen Semester stattfindenden Repetitionsmoduls 6.x wird das erworbene Wissen zur spezifischen Krankheitslehre zusammengeführt und um prozedurales Wissen bezüglich des diagnostischen und therapeutischen Vorgehens bei Vorliegen spezifischer Symptome und Befunde erweitert.

1.2 Prinzipien der Curricularentwicklung

1.2.1 Zielgerichtete Curricularplanung

Laut der Approbationsordnung für Ärzte¹ (ÄApprO) ist das „Ziel der ärztlichen Ausbildung [...] der wissenschaftlich und praktisch in der Medizin ausgebildete Arzt, der zur eigenverantwortlichen und selbstständigen ärztlichen Berufsausbildung, zur Weiterbildung und zu ständiger Fortbildung befähigt ist.“ (Güntert 2003).

Um diesem Anspruch gerecht zu werden, bedarf es einer zielgerichteten Curricularplanung an medizinischen Fakultäten. Eine mögliche Herangehensweise bietet der von Kern et al. (1998) entwickelte sechsstufige Zyklus zur Curricularentwicklung. Basierend auf diesem Ansatz finden in einem ersten Schritt die Identifikation eines Problems und eine allgemeine Bedarfsanalyse statt (Schritt eins). Anschließend wird das zuvor definierte Problem auf eine bestimmte Zielgruppe bezogen, welche spezifische individuelle Bedürfnisse haben kann (Schritt zwei). Im Anschluss an die zielgruppenorientierte Formulierung von Lernzielen (Schritt drei) können Inhalte des Curriculums selektiert und passende Ausbildungsmethoden gewählt werden (Schritt vier). Es folgen Implementierung (Schritt fünf) sowie Evaluation (Schritt sechs) und dynamische Verbesserung des Curriculums.

Lernziele definieren den angestrebten Zugewinn an Wissen und Kompetenzen am Ende eines Lernprozesses. Kompetenzen werden im Sinne der Erziehungswissenschaften allgemein als kognitive, praktische, motivationale und soziale Fähigkeiten zur Lösung von Problemen verstanden (Weinert 2014). Bezüglich der von Bloom et al. (1984) formulierten Lernzieltaxonomie können für Lernziele folgende drei Dimensionen unterschieden werden: Kognitive (Wissen), affektive (Einstellungen) und psychomotorische (praktische Fähigkeiten) Kompetenzen (Kern et al. 1998). Kognitive beziehen sich auf Wissensinhalte (lat. cognoscere = erkennen); affektive Lernziele umfassen Wertvorstellungen, Einschätzungen und Gefühle (lat. afficere = ergriffen sein); psychomotorische Lernziele beinhalten Fertigkeiten des praktischen Handelns (lat. movere = bewegen). In aufeinander aufbauenden Taxonomiestufen wird zwischen den Kompetenzniveaus Kenntnis (niedrigste Stufe), Verständnis, Anwendung, Analyse, Synthese und Evaluation (höchste Stufe) unterschieden (Bloom 1984).

¹ Aus Gründen der besseren Lesbarkeit beschränkt sich diese Arbeit auf die Verwendung männlicher Substantive (u. a. Arzt, Dozent, Lehrender, Spieler) bzw. Pronomina, bezieht sich damit jedoch grundsätzlich in gleichem Maße auch auf weibliche Personen (u. a. Ärztin, Dozentin, Lehrende, Spielerin).

Mittels Erfolgskontrollen sollen Studierende die in Lernzielen vorgegebenen Kompetenzen demonstrieren. Hierzu müssen Prüfungsformate gewählt werden, die im Sinne des „*constructive alignment*“ (Biggs 1996) auf das Lernziel und seine angestrebte Kompetenzebene (siehe auch Kapitel 1.5.2) ausgerichtet sind (Shumway und Harden 2003).

1.2.2 ‚*Outcome-based education*‘

In der medizindidaktischen Forschung herrscht Konsens darüber, dass Lernziele eine Schlüsselposition in der Curricularplanung einnehmen sollen (Harden 1999; Shumway und Harden 2003). „*Consideration of the outcomes should be the basis for curriculum development and evaluation*“, beschreibt beispielsweise Harden (1999) die zentrale Rolle angestrebter Lernziele in der Curricularentwicklung. Dieser Ansatz wird bei dem Paradigma ‚*outcome-based education*‘ verfolgt. Demnach orientieren sich alle Entscheidungen bezüglich der Curricularplanung, insbesondere der Lehrinhalte, der Ausbildungsmethoden und der Prüfungsformate, an zuvor definierten Lernzielen (Harden 1999). Vorbild dieser Ideologie war Spady (1988), der seinen Ansatz einer ergebnisorientierten Lehre mit dem Terminus „*organizing for results*“ zusammenfasste. Spady fokussiert als ersten Schritt dieses Ansatzes die klare Definition von kognitiven, affektiven und psychomotorischen Lernzielen, welche am Ende des Gesamtcurriculums erreicht sein sollen. Ausgehend von diesen Endergebnissen werden zunächst Lernziele für die verschiedenen thematischen Abschnitte des Gesamtcurriculums definiert, um eine kohärent aufeinander abgestimmte Gestaltung der Lehrinhalte zu gewährleisten. Anschließend folgt die strukturierte Formulierung von Lernzielen für einzelne Kurse und schließlich für einzelne Lehrveranstaltungen, um auf differenzierte Weise individuelle Lernerfahrungen zu generieren (Spady 1988). Die gesamte Gestaltung des Curriculums und die Auswahl geeigneter Lehr- und Prüfungsformate unterliegt infolgedessen stets der Orientierung an den angestrebten, stufenweise definierten Lernzielen (Shumway und Harden 2003). So kann der Erwerb spezifischer Kompetenzen gewährleistet und eine größtmögliche Kongruenz zwischen Lernziel, Lehrformat und Prüfungsformat erzielt werden.

Voraussetzung für die Anwendung von ‚*outcome-based education*‘ in der medizinischen Lehre mit dem Ziel der Entwicklung einer kompetenten Arztpersönlichkeit ist die Identifikation von Lernzielen, d. h. von im Laufe des Studiums zu generierenden ärztlichen Kompetenzen, sowie deren transparente Darlegung und deutliche Kommunikation unter allen Beteiligten, einschließlich der Studierenden und der Lehrenden (Harden 1999). Ärztliche Kompetenz als Gesamtkonzept wird nach Epstein und Hundert wie folgt definiert: „*professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions,*

values, and reflection in daily practice for the benefit of the individual and community being served“ (Epstein und Hundert 2002).

1.2.3 Göttinger Lernzielkatalog

An der Medizinischen Fakultät Göttingen wurde für die Lehre im klinischen Studienabschnitt der Göttinger Lernzielkatalog (Göttinger Lernzielkatalog 2008) entwickelt. Im Zuge dessen fand eine Reflexion des Göttinger Curriculums statt, um eine optimale Kongruenz zwischen Lernziel, Lehrmethode und Prüfungsform herzustellen. Gleichzeitig bietet der Göttinger Lernzielkatalog einen hohen Grad an Transparenz gegenüber Lehrenden, Studierenden und der Öffentlichkeit, indem alle prüfungsrelevanten Lehr- und Lerninhalte dargestellt werden. Der Lernzielkatalog umfasst vier sich gegenseitig ergänzende Teilkataloge. Während im ersten Teil Symptome und Befunde aufgelistet werden, umfasst der zweite Teil Gesundheitsstörungen, welche am Ende des Studiums auf verschiedenen Kompetenzniveaus (1-3, wobei „1“ eine vertiefte Kenntnis der Ätiologie, Pathophysiologie, Klinik, Diagnostik und Therapie erfordert) beherrscht werden sollen. In Tabelle 1 sind beispielhaft einige Gesundheitsstörungen mit dem angestrebten Kompetenzniveau dargestellt.

Tabelle 1: Modifiziert entnommen aus dem Göttinger Lernzielkatalog, Teil 2:
Gesundheitsstörungen

Bezeichnung der Gesundheitsstörung	Kompetenz-niveau 1	Kompetenz-niveau 2	Kompetenz-niveau 3
Herzinsuffizienz	KN 1		
COPD	KN 1		
Urolithiasis		KN 2	
Akutes Nierenversagen		KN 2	
Fibrose und Zirrhose der Leber	KN 1		
Divertikulose/Divertikulitis des Darmes		KN 2	
Infektiöse Mononukleose		KN 2	
Akute Leukämie		KN 2	

Abkürzungen: KN 1 = Kompetenzniveau 1; KN 2 = Kompetenzniveau 2

In Teil drei des Katalogs werden erwünschte ethische, rechtliche und wissenschaftliche Kompetenzen benannt, die im vierten Teil durch klinisch-praktische Basisfähigkeiten ergänzt werden.

1.3 ‚Clinical reasoning‘

Eine der wichtigsten und zugleich anspruchsvollsten ärztlichen Kompetenzen stellt die Fähigkeit zu korrekter Diagnosefindung und Therapieeinleitung dar. Die zugrundeliegenden, komplexen Denkprozesse werden als ‚clinical reasoning‘ bezeichnet.

‚Clinical reasoning‘ umfasst unter anderem die Nutzung vielfältiger Strategien zur Generierung, Prüfung und Verifizierung von Diagnosen, das Abwägen von Vorteilen und Risiken diagnostischer und therapeutischer Maßnahmen, sowie nicht zuletzt das Beurteilen von Untersuchungsergebnissen und ihrer prognostischen Relevanz (Kassirer 2010). Diese Kompetenz wird als Kernstück professionellen ärztlichen Handelns (Epstein und Hundert 2002; Norman 2005; Croskerry 2009; Kassirer 2010) angesehen und ihr Erwerb sollte essentieller Bestandteil der medizinischen Lehre sein (Eva 2005; Kassirer 2010).

1.3.1 Zwei-Prozess-Theorien

Grundsätzlich liegen dem ‚clinical reasoning‘ zwei verschiedene Prozesse zugrunde. Hierbei handelt es sich um den intuitiven Ansatz und den analytischen Ansatz (Croskerry 2009). Wissenschaftler haben für diese unterschiedlichen Prozesse zwei verschiedene Systeme des Gehirns postuliert: „System eins“ (intuitiv) und „System zwei“ (analytisch) (Stanovich 1999; Stanovich 2004). Kahneman (2002) beschrieb diese beiden Systeme in seiner Nobelpreisrede wie folgt: „*The operations of System 1 are fast, automatic, effortless, associative, and difficult to control or modify. The operations of System 2 are slower, serial, effortful, and deliberately controlled; they are also relatively flexible and potentially rule-governed*“. Zentraler Aspekt des „System zwei“ (analytisch) ist außerdem seine Korrelation mit Intelligenz und der limitierten Kapazität des Arbeitsgedächtnisses, wohingegen „System eins“ (intuitiv) kaum Gebrauch vom Arbeitsgedächtnis macht (Evans und Stanovich 2013). Intuitive Prozesse sind durch Autonomie charakterisiert und greifen sowohl auf angeborene Methoden und Schemata als auch auf erlernte Strategien zurück, die bis zum Automatismus trainiert wurden (Evans und Stanovich 2013). Somit stützen sich intuitive Prozesse stark auf Erfahrungen und werden angestoßen, wenn bestimmte Muster in gegebenen Informationen wiedererkannt werden (Croskerry 2009). Derartige Prozesse verlaufen schnell, parallel und automatisch; erst das finale Produkt rückt ins Bewusstsein (Evans 2003). Im Gegensatz dazu bestehen analytische Denkprozesse aus hypothetischen Gedankenexperimenten, welche sich an logischen Regeln orientieren (Frankish und Evans 2009) und zu rationalen Entscheidungen führen (Evans und Stanovich 2013). Der Großteil aller Verhaltensweisen wird durch intuitive, unbewusste Prozesse bestimmt. Sollte die Komplexität einer Situation analytische Denkprozesse,

Reflexion und den Zugriff auf das Arbeitsgedächtnis erfordern, greift allerdings das analytische „System zwei“ ein (Evans und Stanovich 2013).

1.3.2 Der diagnostische Entscheidungsprozess

Essentieller Bestandteil des ‚*clinical reasoning*‘ in der adäquaten Patientenversorgung ist der Prozess der Diagnosefindung. Die folgende Darstellung dieses Prozesses bezieht sich auf die Publikation von Croskerry (2009). Den ersten Schritt dieses diagnostischen Entscheidungsprozesses stellt demnach die Präsentation der Patienteninformationen und Krankheitszeichen dar. Derartige Informationen werden gewöhnlich im persönlichen Kontakt mit dem Patienten gewonnen, können aber, im Hinblick auf die medizinische Ausbildung, auch durch schriftliche Beschreibungen oder virtuelle Patienten gegeben werden.

Die Daten werden im zweiten Schritt auf pathognomonische Krankheitszeichen untersucht. Im Falle der Erkennung bekannter Muster wird schnell und automatisch eine Antwort des intuitiven „System eins“ angestoßen. Sollten die Krankheitszeichen hingegen nicht eindeutig oder der Entscheidungsträger unsicher sein, wird das analytische „System zwei“ eingeschaltet. Durch dieses erfolgt eine objektive und systematische Auseinandersetzung mit den gegebenen Informationen mit dem Ziel logischer Schlussfolgerungen. Es werden Hypothesen geprüft sowie Differentialdiagnosen erwogen und untersucht, um eine kohärente klinische Geschichte zu entwickeln (Kassirer 2010). Durch die Repetition analytischer Prozesse werden neue Muster erlernt, welche bei der erneuten Konfrontation mit entsprechenden Inhalten erkannt werden und schließlich eine intuitive Antwort („System eins“) auslösen. Beide Systeme sind in der Lage, sich über das jeweils andere System hinwegzusetzen. So wird die Fähigkeit zur Metakognition durch das „System zwei“ ermöglicht, indem es alle schnellen, intuitiven Prozesse überwacht und das „System eins“ zur Neubewertung von Informationen zwingen kann. Andererseits kann übermäßiges Selbstvertrauen des Entscheidungsträgers zur Verfolgung intuitiver Ansätze führen, auch wenn rationale Schlussfolgerungen ein anderes Verhalten bedingen würden.

Die zwei Systeme sollten demnach nicht dichotom, sondern additiv gesehen werden; sowohl intuitive als auch analytische Prozesse leisten einen Beitrag zur finalen Entscheidung (Eva 2005). Während in der initialen Phase eines neuen Patientenfalls intuitive Prozesse zur Generierung von Hypothesen dominieren (Eva 2005; Kassirer 2010), stehen in der anschließenden Prüfung dieser Hypothesen analytische Prozesse im Fokus (Eva 2005). Darüber hinaus unterliegt der Prozess der Entscheidungsfindung einer enormen

Kontextspezifität bezüglich situationsbedingter und persönlicher Faktoren (Eva 2005), wodurch die bevorzugte Anwendung intuitiver bzw. analytischer Prozesse forciert werden kann.

1.3.3 Zentrale Elemente des ‚*clinical reasoning*‘

Nachdem der diagnostische Entscheidungsprozess nun erläutert wurde, sollen im folgenden Abschnitt die zentralen Elemente des ‚*clinical reasoning*‘ zusammengeführt werden.

In der Annäherung an einen neuen Patientenfall steht zunächst die Datenerfassung im Vordergrund (Bowen 2006). Dieser Aspekt berücksichtigt unter anderem die Erhebung der Patientengeschichte, die Durchführung einer körperlichen Untersuchung oder die Interpretation der Ergebnisse aus Laboruntersuchung und Bildgebung (Bowen 2006). Anschließend steht die Transformation der patientenspezifischen Informationen in abstrakte Deskriptoren (sog. ‚*semantic qualifiers*‘) im Fokus, um Hypothesen generieren und Daten vergleichbar machen zu können. So wird beispielsweise die anamnestische Angabe „Ich hatte diese Probleme bereits früher schon einmal“ zu „wiederkehrend“ (Bowen 2006). Auf diese Weise wird der Abgleich mit Erinnerungen und bekannten Mustern erleichtert, um gespeichertes klinisches Wissen abzurufen und dem Gedankenprozess zugänglich zu machen (Bowen 2006). Laut Bowen (2006) erfolgt der Abruf klinischen Wissens in Form von ‚*illness skripts*‘, welche eine Fülle an klinisch relevanten Informationen enthalten. Durch den analytischen Abgleich der gewonnenen Patientendaten und der generierten Hypothesen mit den Charakteristika verschiedener ‚*illness skripts*‘ können diagnostische Hypothesen verfeinert und verworfen oder bestätigt werden. Auf dem Weg zum Experten nimmt, mit zunehmender Erweiterung des klinischen Erfahrungsschatzes, die Geschwindigkeit dieser Analyse zu: Krankheitszeichen werden immer schneller in gelernten Mustern und ‚*illness skripts*‘ wiedererkannt, bis dieser Vorgang schließlich intuitiv erfolgt. Mit Abruf des ‚*illness skripts*‘ kann außerdem ein Plan zum therapeutischen Management des Patienten entworfen werden. Hierbei muss unter anderem zwischen ‚*watchful waiting*‘ (d. h. Abwarten und Beobachten) und der Notwendigkeit zur sofortigen Intervention entschieden werden, stets auf Basis eines ausgewogenen Nutzen-Risiko-Verhältnisses und in Kenntnis der Behandlungsgrenzen (Kassirer 2010).

1.3.4 Konsequenzen für die medizinische Lehre

‚*Clinical reasoning*‘ stellt eine zentrale Komponente ärztlicher Kompetenz dar und sollte sowohl gelehrt als auch getestet werden (Norman 2005). In den vorherigen Abschnitten wurde ausgeführt, dass intuitive und analytische Prozesse additiv zum Lösen klinischer

Probleme genutzt werden sollten (Eva 2005; Ark et al. 2006). Während in vielen Fällen das Erkennen von Mustern klinischer Symptome mit dem Abruf des korrekten ‚*illness skript*‘ eine gute Strategie darstellt, müssen in anderen Fällen aufgrund von Kontext oder Komplexität diagnostische Algorithmen angewandt oder wissenschaftliche Grundlagen der Pathophysiologie berücksichtigt werden (Eva 2005). Aus diesem Grund sollten Studierenden vielfältige Strategien zum Lösen klinischer Probleme vermittelt werden (Eva 2005). Zunächst muss Grundlagenwissen angeeignet werden, eine fundamentale Voraussetzung für erfolgreiches ‚*clinical reasoning*‘ (Kassirer 2010). In einem anschließenden Studienabschnitt sollten Studierende animiert werden, ihr Faktenwissen anzuwenden und zu reorganisieren (Norman et al. 2016). Nur durch das Schaffen von Verbindungen zwischen Faktenwissen einerseits und klinischen Situationen andererseits können Studierende Zusammenhänge zwischen Krankheitsbildern und dem im Gedächtnis gespeicherten Wissen herstellen (Bowen 2006). Von besonderer Bedeutung ist in diesem Sinne das Kennenlernen von Fallbeispielen (Eva 2005; Kassirer 2010) zur Entwicklung einer mentalen Datenbank an ‚*illness skripts*‘, auf deren Basis nicht-analytische Prozesse stattfinden können (Eva 2005). Anhand von Fallbeispielen können Studierende befähigt werden, bei der Entscheidungsfindung situationsbedingt und flexibel sowohl auf analytische Prozesse als auch auf intuitive Mustererkennung zurückzugreifen (Bowen 2006). Die Fälle müssen so ausgewählt sein, dass sie die relevanten Aspekte (vgl. Kapitel 1.3.3) eines erfolgreichen ‚*clinical reasoning*‘ abbilden und fordern (Kassirer 2010). Außerdem sollten Fallbeispiele verschiedener Fachbereiche und diagnostischer Kategorien gemischt präsentiert werden, um Flexibilität in den Denkprozessen zu trainieren (Eva 2005).

1.4 Lehrformen

In der Kognitionspsychologie wird Lernen als einen aktiver, konstruktiver und zielgerichteter Prozess angesehen, welcher von den geistigen Aktivitäten des Lernenden abhängt (Shuell 1986). Lernen ist ein Prozess der aktiven Informationsverarbeitung, nicht der Informationsvermittlung (van der Vleuten und Driessen 2014). Gute Lehre fördert den Prozess der Informationsverarbeitung durch die Lernenden (van der Vleuten und Driessen 2014), indem eine Verschiebung von traditionellem ‚Lehrer-zentriertem‘ zu einem ‚Studierenden-zentrierten‘ Unterricht erfolgt (Biggs 1996; Towle und Cottrell 1996; Spencer und Jordan 1999). Im Gegensatz zum ‚oberflächlichen Lernen‘ (welches zur Reproduktion auswendig gelernten Wissens befähigen kann) können so Lernsituationen geschaffen werden, welche Studierende zu ‚tiefem Lernen‘, d. h. der Konstruktion von Wissen (Biggs 1996) und dem aktivem Verständnis (Spencer und Jordan 1999), ermutigen.

Im Hinblick auf eine Kongruenz zwischen Lernziel und Lehrformat (s. o.) sind traditionelle ‚Lehrer-zentrierte‘ und neuartige ‚Studierenden-zentrierte‘ Lehrformate in unterschiedlichem Maß geeignet, kognitive, affektive und psychomotorische Kompetenzen zu vermitteln und auf dem geforderten Kompetenzniveau zu erreichen. Im Folgenden werden einige Lehrformate mit ihren kongruenten Lernzielen und Kompetenzebenen vorgestellt.

1.4.1 Die Vorlesung

Vorlesungen stellen ein traditionsreiches und bis heute dominierendes Lehrformat an Hochschulen dar (Watts und Becker 2008; van der Vleuten und Driessen 2014). In den meisten Fällen hält eine Lehrkraft einen Vortrag im klassischen Frontalunterricht und vermittelt einer Vielzahl von Studierenden gleichzeitig Wissen. Die hauptsächlichen Tätigkeiten der Studierenden während einer Vorlesung sind Zuhören, Interpretieren, Verstehen und das Notieren von Stichpunkten (Biggs 1996). Die Lernenden nehmen also während der Vorlesung in erster Linie eine passiven Rolle ein, sie konsumieren Informationen, ohne diese zu verarbeiten (Biggs 1996; van der Vleuten und Driessen 2014) und akzeptieren Fakten häufig, ohne sie zu hinterfragen (Biggs 1996). Dementsprechend stellen Vorlesungen ein geeignetes Lehrformat für die Vorstellung von Faktenwissen und die Vermittlung von Grundlagenwissen zu Krankheitsbildern dar. Für die Vermittlung von Lernzielen höherer Ebenen müssen jedoch andere Lehrformate gewählt werden.

1.4.2 Das Seminar

Die ÄApprO (Güntert 2003) sieht als Unterrichtsveranstaltungen neben Vorlesungen insbesondere Seminare vor, in denen die Lerninhalte „vertiefend, anwendungs- und gegenstandsbezogen“ bearbeitet werden sollen. Es handelt sich um einen Kleingruppenunterricht, in dem einem Seminarleiter maximal 20 Teilnehmer zugeteilt sind. Zu den wesentlichen Zielen von Seminaren zählt die Elaboration und Vertiefung (vor-) klinischer Zusammenhänge. Insgesamt müssen Seminare im Umfang von mindestens 98 Stunden als integrierte Veranstaltungen unter Einbezug klinischer Fächer sowie im Umfang von mindestens 56 Stunden als Veranstaltungen mit klinischem Bezug angeboten werden. Bei Seminaren handelt sich um ‚Studierenden-zentrierten‘ Unterricht, in welchem die aktive Mitarbeit der Teilnehmer gefordert wird (Jaques 2003) und die Studierenden eigene Beiträge zu erbringen haben (Güntert 2003). Die Studierenden können in dieser Lernumgebung fächerübergreifende Probleme diskutieren, Meinungen anderer Studierender in Frage stellen und Missverständnisse klären (Biggs 1996). Darüber hinaus werden auch soziale Kompetenzen wie das Zuhören, die Präsentation von Ideen, das Arbeiten im Team (Jaques

2003) oder Kommunikationsfähigkeit und Respekt trainiert (Wood 2003). Seminare sind somit geeignet, insbesondere affektive Lernziele zu vermitteln und auf einem höheren Kompetenzniveau (prozedurales Wissen/Handlungs- und Begründungswissen) zu erreichen. Die Effektivität hängt jedoch unter anderem von der Gruppengröße und -konstellation sowie der Kompetenz des Seminarleiters ab (Biggs 1996). Beispielsweise verfallen die Seminarleiter häufig in einen ‚Lehrer-zentrierten‘, Frontalunterricht zurück oder einzelne Studierende dominieren bzw. blockieren die Diskussionsrunde (Jaques 2003). Darüber hinaus sind Seminare schlecht geeignet für heterogene Studierendengruppen, da sie kein individuelles Lerntempo zulassen (Jaques 2003).

1.4.3 Das Problemorientierte Lernen

Bei ‚Problemorientiertem Lernen‘ handelt es sich um eine Ausgestaltung des Seminars, in welchem Studierende in Kleingruppen klinische Problemstellungen bearbeiten, um einen Wissens- und Verständniszuwachs zu erzielen (Wood 2003). Dieses Lehrformat wurde im Jahr 1969 von Neufeld und Barrows (1974) erstmalig an der McMaster Universität in Kanada (Hamilton, Ontario) und in den folgenden Jahren an zahlreichen weiteren internationalen und deutschen Universitäten implementiert (Spencer und Jordan 1999). Der Grundgedanke ist, dass der Lernprozess mit einem klinischen Problem bzw. einem klinischen Fall beginnt (Towle und Cottrell 1996), welcher als Stimulus für aktives, selbstständiges Lernen (Davis und Harden 1999) wirkt. Die Kleingruppen bestehen aus etwa fünf bis zehn Studierenden und einem Tutor, dessen Aufgabe es ist, die Studierenden in ihrem selbstbestimmten Lernprozess zu begleiten und zu unterstützen (Towle und Cottrell 1996). Üblicherweise erfolgt der Ablauf des Unterrichts in sieben Schritten nach dem Maastricht ‚seven jump‘-Programm: 1) Fallvorstellung und Klärung von Verständnisfragen; 2) Problemdefinition; 3) Problemanalyse; 4) Hypothesengenerierung; 5) Lernzieldefinition; 6) Selbststudium und Recherche; 7) Berichterstattung, Synthese und Anwendung des neu erworbenen Wissens auf das Problem (Spencer und Jordan 1999). Allerdings gibt es vielfältige Ausgestaltungsmöglichkeiten und Varianten von problembasiertem Lernen (Davis und Harden 1999; Reusser 2005). Im Allgemeinen werden hierzu alle Lehr- bzw. Lernsituationen gezählt, in welchen Studierende zur selbstständigen Lösung von realen oder realitätsnahen Problemen angeregt werden, ohne dass die oben genannten sieben Schritte vollständig durchlaufen werden müssen (Reusser 2005).

Eine Variante von problembasiertem Lernen stellt das an der UMG im Rahmen der Pflichtlehre des sechsten klinischen Semesters implementierte POLEMA-Konzept dar (siehe Kapitel 2.3.1). POLEMA steht für „Problem orientiertes Lernen – elektronische Medien

assistiert“ und ist eine Lehrform, in der Studierende in Kleingruppen unter Anleitung studentischer Tutoren echte Patientenfälle besprechen. Basierend auf einem vorgestellten Krankheitsbild sollen die Studierenden Differentialdiagnosen nennen und selbstständig sinnvolle diagnostische und therapeutische Vorgehensweisen entwickeln. Der Ablauf des Lehrformats orientiert sich grundsätzlich am Maastricht ‚seven jump‘-Programm (s. o.): Der Unterricht beginnt mit der Vorstellung eines klinischen Falls, welcher in Kleingruppen diskutiert wird. Anschließend betreiben die Studierenden Selbststudium und Recherche, um ihre zuvor definierten Lernziele zu erreichen und den klinischen Fall zu lösen. In einer abschließenden Besprechung im Plenum präsentieren die Gruppen ihre Ergebnisse. Jedoch weicht das POLEMA-Konzept insofern vom originalen Konzept ab, als dass den Studierenden spezifische Vorgaben zur Gestaltung und zum Ablauf der Gruppenarbeitstermine gemacht werden. So sollen zum Beispiel zunächst Differentialdiagnosen und anschließend diagnostische und therapeutische Schritte benannt werden.

Problembasiertes Lernen motiviert Studierende, selbstständig individuelle Lernziele und Fragen zu formulieren (Wood 2003) und durch das Selbststudium und die Diskussion mit den Gruppenteilnehmern zu erreichen bzw. zu beantworten. Auf diese Weise wird ‚tiefes Lernen‘ gefördert (Davis und Harden 1999; Wood 2003) und eine stimulierende, herausfordernde Lernumgebung geschaffen (Wood 2003). Die Auseinandersetzung mit Fallbeispielen fördert darüber hinaus den Erwerb von ‚*illness skripts*‘ (s. o.) (Spencer und Jordan 1999). Problembasiertes Lernen ist damit gut zur Vermittlung von Lernzielen höherer Ebenen und auch zum Erwerb einer eigenständigen Lernkompetenz geeignet. Allerdings handelt es sich um ein sehr ressourcenintensives Lehrformat (Des Marchais 1993; Finucane et al. 1998; Spencer und Jordan 1999; Wood 2003; Srinivasan et al. 2007).

1.4.4 Der Unterricht am Krankenbett

Der „Unterricht am Krankenbett“ stellt ein aktives Kleingruppenlehrformat im realen Kontext unter Anwesenheit echter Patienten dar (Nair et al. 1998; Aldeen und Gisondi 2006; Salam et al. 2011). In der ÄApprO wird dem „Unterricht am Krankenbett“ ein hoher Stellenwert zugeschrieben: Bei einer bindenden Gesamtstundenzahl von 476 Stunden soll die Unterrichtszeit je zur Hälfte auf die Patientendemonstration (in einer Gruppe von höchstens sechs Studierenden) und die Patientenuntersuchung (in einer Gruppe von höchstens drei Studierenden) entfallen (Güntert 2003).

Neben dem Erwerb von Faktenwissen bezieht sich dieses Lehrformat vor allem auf psychomotorische und affektive Lernziele. Insbesondere werden diagnostische (z. B. Anamneseerhebung; körperliche Untersuchung) und administrative Fähigkeiten (z. B. Zeitmanagement) trainiert (Nair et al. 1998). Studierende können die Interaktion mit echten Patienten üben und eine Arzt-Patienten-Beziehung aufbauen (Qureshi und Maxwell 2012; Peters und ten Cate 2014). Darüber hinaus können in dieser Lernumgebung Professionalität (LaCombe 1997; Aldeen und Gisondi 2006; Salam et al. 2011), Grundsätze der ärztlichen Berufsethik (Peters und ten Cate 2014) und Humanismus (Aldeen und Gisondi 2006) vermittelt werden.

Trotz der Vermittlung relevanter Kompetenzen ist der Einsatz von „Unterricht am Krankenbett“ stark rückläufig, was sich in erster Linie durch einen betriebswirtschaftlich bedingten, deutlich gestiegenen Patientendurchlauf in Krankenhäusern erklären lässt (Peters und ten Cate 2014). Die kürzere Patientenverweildauer bedingt nicht nur einen höheren Arbeitsaufwand und stärkeren Zeitdruck für behandelnde Ärzte, sondern reduziert auch die Verfügbarkeit von Patienten für den „Unterricht am Krankenbett“ (Nair et al. 1998; Qureshi und Maxwell 2012; Peters und ten Cate 2014). Es handelt sich um ein sehr zeit- und personalintensives Unterrichtsformat (Aldeen und Gisondi 2006), dessen Umsetzung weitere Herausforderungen birgt. Hierzu zählen unter anderem der unruhige Stationsalltag, welcher die Situation der fehlenden Verfügbarkeit von Patienten verschärft (z. B. aufgrund von Untersuchungen) und das Erzeugen einer konzentrierten Lernumgebung beeinträchtigt (z. B. aufgrund von Telefonaten oder Besuchern) (Nair et al. 1998). Darüber hinaus müssen Lehrende eine heterogene Gruppe von Studierenden mit unterschiedlichen Wissensständen gleichzeitig unterrichten (Kroenke 1992). Nicht zuletzt können die Unterrichtsinhalte nicht standardisiert werden, da sie von den zur Verfügung stehenden Patienten und ihren Krankheitsbildern abhängen (Raupach et al. 2016). Aus diesem Grund kann nicht sichergestellt werden, dass alle Studierenden gegenüber den gleichen Inhalten exponiert werden. Eine qualitative Studie von Gierk und Harendza (2012) ergab, dass bei der Auswahl von Patienten durch die Lehrenden neben den Krankheitsbildern auch strukturelle und bio-psycho-soziale Faktoren berücksichtigt werden. So zählen zu den Ausschlusskriterien beispielsweise ein schlechter Gesundheitszustand oder ansteckende Krankheiten, wodurch es zu einer nicht-repräsentativen Patientenauswahl und Selektionsverzerrung kommen kann. Ferner ergeben sich im Rahmen des „Unterrichts am Krankenbett“ nur selten Unterrichtsgelegenheiten zur Diskussion akuter Notfälle. Zudem erscheint eine patientenzentrierte Lehre in Notfallsituationen auch ethisch nicht immer vertretbar, da eine Verpflichtung gegenüber dem Patienten besteht, durch eine optimale medizinische

Versorgung größtmögliche Sicherheit zu gewährleisten (Ziv et al. 2003). Um auch die Versorgung akut erkrankter oder infektiöser Patienten und die Entscheidungsfindung in Akutsituationen zu trainieren, müssen demnach andere Lehrformate, wie zum Beispiel Simulationen, gewählt werden (Gierk und Harendza 2012).

1.4.5 Simulationen

In der medizinischen Ausbildung werden verschiedene Arten von Simulationen eingesetzt. Hierbei handelt es sich unter anderem um Rollenspiele, Simulationspatienten, Simulatorpuppen und computerbasierte Simulationen (Lane et al. 2001). Rollenspiele, in denen Studierende die Rolle von Ärzten bzw. Patienten imitieren, und Simulationspatienten, die ausgebildet wurden, um Erkrankungen und Symptome standardisiert darzustellen (Barrows 1993), eignen sich insbesondere zum Ausbau kommunikativer und interpersoneller Kompetenzen und zum Training körperlicher Untersuchungsmethoden (Barrows 1993). Mithilfe von Simulatorpuppen können zudem pharmakologische Aus- und Wechselwirkungen sowie komplexe apparative Untersuchungen und Behandlungsmaßnahmen, beispielsweise die kardiologische Auskultation oder die endotracheale Intubation, vermittelt werden (Lane et al. 2001). Computerbasierte Simulationen verknüpfen die Vermittlung von Lerninhalten mit interaktiven Elementen (Lane et al. 2001). Eine innovative Art computerbasierter Simulationen stellen Lernspiele, sog. ‚*Serious Games*‘, dar.

Im Rahmen der hier vorliegenden Arbeit wurde das ‚*Serious Game*‘ EMERGE in die curriculare Lehre der UMG implementiert (Näheres hierzu siehe Kapitel 2.3.2). Um ein besseres Verständnis dieser Intervention zu ermöglichen, folgen im nächsten Kapitel nähere Ausführungen zum Lehrformat ‚*Serious Game*‘ im Allgemeinen und zu seinem Einsatz in der medizinischen Lehre.

1.4.6 ‚*Serious Games*‘

Laut Friedmann (2000) ist die medizinische Ausbildung „*stuck in time, space and content*“, denn a) gute Lehrveranstaltungen erfordern die gleichzeitige Interaktion von Lernenden und Lehrenden, b) die Wissensvermittlung ist räumlich an spezifische Orte und Unterrichtsräume gebunden und c) Unterrichtsinhalte können nur unzureichend durch Studierende und Lehrende kontrolliert werden. Dies steht in Diskrepanz zum Lernverhalten der heutigen, in das digitale Zeitalter hineingeborenen Studierendengeneration (sog. ‚*Digital Natives*‘), welche Technologien wie das Internet und Computerspiele in ihr Leben integriert haben (Prensky 2001). Sie können sich jederzeit Informationen in den digitalen Medien beschaffen und

Wissen aktiv konstruieren (Barnes et al. 2007). Es resultiert eine zeitliche, personelle und örtliche Unabhängigkeit im Prozess des Wissenserwerbs und eine Forderung nach autonomem Lernen von Seiten der Studierenden (Barnes et al. 2007). Durch den angemessenen Einsatz technologiebasierter Unterrichtsmethoden in der medizinischen Ausbildung kann die Lehre weiterentwickelt und an die Bedürfnisse der modernen Studierendengeneration angepasst werden (Friedman 2000). Der Ausbau digitaler Infrastruktur und die zunehmende Verfügbarkeit von Computermedien in Lehr- und Lerneinrichtungen ermöglicht die Entwicklung von computerbasierten 3D-Simulationen (dreidimensionale Simulationen), in welchen Studierende mit didaktisch aufbereiteten Problemen konfrontiert werden. Eine Umfrage von Kron et al. (2010) ergab, dass Medizinstudierende den Einsatz neuer Medien in der Lehre positiv bewerten (98%) und insbesondere im Einsatz von Computerspielen einen bildungsspezifischen Nutzen sehen (80%).

1.4.6.1 Begriffseinordnung

Der Oberbegriff für alle Varianten technologiebasierter Lehr- und Lernmethoden ist ‚*E-Learning*‘ (englisch ‚*Electronic Learning*‘ = elektronisches Lernen). ‚*E-Learning*‘ ist ein Sammelbegriff für unterschiedliche Technologien und Unterrichtsmaterialien wie u. a. fallorientiertes Lernen, Hypermedia, Simulationen (Ruiz et al. 2006) oder ‚*Serious Games*‘. Letztere sind Computerspiele, die auf den Erwerb von Wissen und die Entwicklung von Fähigkeiten (Vivekananda-Schmidt und Psycholb 2011; Girard et al. 2013) sowie auf die Verbesserung von Fachkompetenzen (Graafland et al. 2012) gerichtet sind. Es existiert eine große Vielfalt an ‚*Serious Games*‘, sowohl bezüglich des Genres (Strategiespiel, Abenteuerspiel etc.) als auch bezüglich der vermittelten Kompetenzen (Girard et al. 2013). Ein verbreitetes Spielgenre, vor allem in der medizinischen Ausbildung, stellen Simulationsspiele dar, welche Lernsituationen in einer virtuellen Umgebung ermöglichen (Kiili 2005). In der Konfrontation mit virtuellen Patienten kann die verantwortliche Rolle bezüglich der Anamneseerhebung, der körperlicher Untersuchung sowie der diagnostischen und therapeutischen Entscheidungen nachgeahmt werden (Association of American Medical Colleges 2007).

Die didaktische Verknüpfung von Varianten des ‚*E-Learning*‘ mit traditionellen Präsenzveranstaltungen wird als ‚*Blended Learning*‘ bezeichnet (Ruiz et al. 2006). Hierbei wird die Integration der Vorteile von asynchronem ‚*E-Learning*‘ und von synchronem Präsenzlernen in einem gemeinsamen Curriculum angestrebt (Garrison und Kanuka 2004). Im Gegensatz zum synchronen Lernen, bei dem die Interaktion und Kommunikation

zwischen Lernenden und Lehrenden zeitgleich zur Wissensvermittlung stattfindet, findet die Interaktion zwischen Lehrenden und Lernenden beim asynchronen Lernen zeitlich und räumlich versetzt statt (Er et al. 2009). Auf diese Weise kann das Lernen an individuelle Lerngeschwindigkeiten und Bedürfnisse angepasst werden. Die Kombination von asynchronem und synchronem Lernen im Rahmen eines ‚*Blended Learning*‘-Formats schafft einerseits Unabhängigkeit und Selbstverantwortung im (asynchronen) Lernprozess der Studierenden, wodurch diese zu komplexen Gedankenprozessen und kritischem Denken angeregt werden, andererseits die Möglichkeit zum direkten Austausch im Rahmen von (synchronen) Präsenzveranstaltungen (Garrison und Kanuka 2004).

1.4.6.2 Das Erreichen von Lernzielen durch ‚*Serious Games*‘

Das „*Input-Process-Outcome Game Model*“ von Garris et al. (2002) beschreibt den Lernprozess, der während des Spielens von ‚*Serious Games*‘ durchlaufen wird. Der Input besteht aus dem ‚*Serious Game*‘ mit der Integration von Lehrinhalten und charakteristischen Spieleigenschaften (z. B. Regeln/Ziele, Herausforderung, Fantasie). Dieser Input löst einen repetitiven Spielzyklus aus den drei Komponenten Verhalten, Feedback und Urteil aus. Jedes Verhalten des Spielers bedingt ein Feedback durch das System. Hieraus zieht der Spieler Rückschlüsse, welche Einfluss auf das Verhalten im nächsten Spielschritt haben. Durch diesen Interaktionszyklus werden beim Spieler unterschiedliche Reaktionen wie Freude, Spaß oder Ehrgeiz hervorgerufen, die zum Weiterspielen motivieren und Lernerfahrungen ermöglichen. Die Reflexion und Analyse der Spielereignisse (das sog. ‚*Debriefing*‘) führt schließlich zum Erreichen von Lernzielen. Laut Connolly et al. (2012) ist „Wissensakquisition“ die durch Lernspiele am häufigsten erreichte Zielgröße. Im Gegensatz dazu werden affektive und motivationale Erfolge vor allem durch Computerspiele ohne didaktische Intention erzielt (Connolly et al. 2012). Issenberg et al. (2005) identifizierten allerdings, dass sich bislang nur wenige Forschungsarbeiten auf Lernziele wie Managementfähigkeiten, klinische Fähigkeiten oder medizinisches Grundlagenwissen konzentrieren.

1.4.6.3 Die Vorteile von ‚*Serious Games*‘

Laut Cave et al. (2007) fühlen sich lediglich 59% der Medizinstudierenden ausreichend auf den klinischen Arbeitsalltag vorbereitet. Dies ist vor allem auf eine mangelnden Exposition gegenüber klinischer Praxis und direkten Patientenkontakten zurückzuführen (Bloice et al. 2014). Diese Exposition ist hingegen Voraussetzung für das Schaffen von logischen Verbindungen zwischen Faktenwissen und Praxiswissen sowie für seine Organisation in Form von ‚*illness skripts*‘ (siehe Kapitel 1.3.3) (Bowen 2006). Der Bedarf kann durch ‚*Serious*

Games‘ angesprochen werden, in welchen Studierende mit virtuellen Patienten und realen klinischen Problemen konfrontiert werden. In diesem Lehrformat können Fähigkeiten und Inhalte zielgerichtet geübt und vielfach wiederholt werden, wodurch laut Ericsson (2004) Expertise erlangt wird. Studierenden wird Zugang zu einer Vielzahl von unterschiedlichen Krankheitsbildern verschafft, was die Entwicklung von ‚*illness skripts*‘ und die Schulung von ‚*clinical reasoning*‘-Fähigkeiten ermöglicht (Bloice et al. 2014). Die interaktive, studierendenzentrierte Lernerfahrung schafft einen starken Lernanreiz (Ruiz et al. 2006) und motiviert zum intensiven Training klinischer Problemlösefähigkeiten (Bryce et al. 1998).

In einer qualitativen Studie von Bryce et al. (1998) schätzten Studierende insbesondere das hohe Maß an Unabhängigkeit und Autonomie im Lernprozess. Erstens ist das Bearbeiten klinischer Fragestellungen weder von der Verfügbarkeit von Dozenten und Patienten (Bryce et al. 1998) noch von Zeit und Ort abhängig (Garrison und Kanuka 2004). Zweitens kann der Lernende den gesamten Verlauf der Patienteninteraktion inklusive Untersuchung, Befundinterpretation, Hypothesengenerierung und Behandlung selbst bestimmen (Bryce et al. 1998) und lernen, Verantwortung für Tätigkeiten und ihre Konsequenzen zu übernehmen (Kaczmarczyk et al. 2015). Ohne Gefahr für reale Patienten können unterschiedliche Lösungswege ausprobiert und Ursache-Wirkungs-Zusammenhänge verstanden werden (Bearman 2003); Fehler sind erlaubt und können ohne Angst vor Schuld- und Haftungsfragen aufgedeckt und korrigiert werden (Friedman 2000; Ziv et al. 2003; Ericsson 2004; Issenberg et al. 2005). Auf diese Weise ermöglichen ‚*Serious Games*‘ eine patientenzentrierte Lehre auch in komplexen oder akut lebensbedrohlichen Situationen und seltenen Krankheitsfällen (Ziv et al. 2003; Issenberg et al. 2005), in denen die Verantwortungsübertragung auf Studierende in der Realität nicht vertretbar ist. Überdies vermittelt der Einsatz von Simulationen zum Schutz von Patienten auch eine ethische Botschaft: Das Patientenwohl hat oberste Priorität und Patienten dürfen nicht als Mittel zum Ausbildungszweck missbraucht werden (Ziv et al. 2003).

Eine weitere Stärke von ‚*Serious Games*‘ stellt die Standardisierung der Lehrinhalte dar (Ruiz et al. 2006), wodurch sichergestellt werden kann, dass alle Studierenden gegenüber den gleichen patientenbezogenen Inhalten exponiert werden. Diese Inhalte können an persönliche Bedürfnisse und Lernstrategien angepasst, im eigenen Tempo erarbeitet (Issenberg et al. 2005; Wahlgren et al. 2006) und beliebig oft wiederholt werden, um Selbstvertrauen und Handlungskompetenz auf einem komfortablen Niveau zu erreichen (Ziv et al. 2003). Da der Computer jeden einzelnen Schritt des Interaktionsprozesses vom Lernenden mit den Lehrinhalten erfasst, kann ein hoch-spezifisches Feedback bezüglich der gezeigten Stärken und Schwächen generiert werden (Friedman 2000). Nicht zuletzt sind

„*Serious Games*“ ein ökonomisches Lehrformat, da zahlreiche Studierende gleichzeitig unterrichtet und auf diese Weise zeitliche und personelle Ressourcen eingespart werden können (Lyon et al. 1991; Dankbaar et al. 2017).

1.4.6.4 Effektivität von „*Serious Games*“

Hinsichtlich der Effektivität von „*Serious Games*“ finden sich in der Literatur bislang wenige empirische Studien (Girard et al. 2013). Die Untersuchungen konzentrieren sich hierbei hauptsächlich auf Akzeptanz und studentischen Lernerfolg. Zuverlässige Rückschlüsse müssen allerdings unter Vorsicht gezogen werden: Einerseits aufgrund der Vielfalt an „*Serious Games*“, andererseits aufgrund der großen Unterschiede zwischen existierenden Studien in Bezug auf Forschungsfragen, wissenschaftliches Vorgehen und Studiendesign (Girard et al. 2013).

Im Allgemeinen hat der Einsatz von „*E-Learning*“-Materialien in der medizinischen Ausbildung positive Effekte auf den Erwerb von Wissen und „*clinical reasoning*“-Fähigkeiten verglichen mit keiner Intervention (Lyon et al. 1991; Carr et al. 1999; Wahlgren et al. 2006). In eine von Cook et al. (2008) veröffentlichte Metaanalyse gingen insgesamt 201 Studien verschiedener Sprachen ein, welche den Einsatz von internetbasierten Lehrformaten in Gesundheitsberufen und ihre Effektivität im Hinblick auf Zufriedenheit und Lernerfolg mit keiner Intervention oder mit dem Einsatz nicht-internetbasierter Lehrformate verglichen. Es zeigte sich, dass der Einsatz internetbasierter Lehrformate sehr effektiv im Vergleich zu keiner Intervention war. Im Vergleich zum Einsatz traditioneller Lehrformate ergaben sich jedoch lediglich kleine Effektgrößen. Hier zeigte sich also eine vergleichbare Effektivität von internetbasierten und traditionellen Lehrformaten. Weitere Studien stufen „*E-Learning*“ im Vergleich mit traditionellem Unterricht in Bereichen wie Wissenserwerb in der Dermatologie (Wahlgren et al. 2006) oder dem Management von Adominalschmerz (Vash et al. 2007), Epistaxis (Carr et al. 1999), Anämie und Brustschmerz (Lyon et al. 1991) als gleichwertig hinsichtlich des Lernerfolgs ein. Ebenso konnte für bestimmte „*Serious Games*“ ein mindestens gleichwertiger Lernerfolg im Vergleich zu traditionellen Lehrformaten identifiziert werden. In einer Studie von Dankbaar et al. (2016a) wurde ein „*Serious Game*“ zur Vermittlung von Wissen und Fähigkeiten hinsichtlich der Stabilisierung von Patienten in Akutsituationen eingesetzt. Studierende zeigten gleiche Lernerfolge bezüglich kognitiver, klinischer und kommunikativer Fähigkeiten wie Studierende, die dieselben Inhalte entweder mithilfe von textbasierten Patientenfällen in einer Gruppe oder mithilfe von Online-Lernmaterialien gelernt hatten. Ähnliche Ergebnisse zur Gleichwertigkeit von bestimmten „*Serious Games*“ und traditionellem Unterricht finden sich unter anderem für den Erwerb von Wissen zur

Patientensicherheit (Dankbaar et al. 2017), für den Erwerb von Wissen und Fähigkeiten zur Insulintherapie (Diehl et al. 2015) und für die Verbesserung von Teamarbeits- und Teamführungskompetenzen (Youngblood et al. 2008).

Für den Einsatz virtueller Patienten ergab sich in einigen Studien sogar eine Überlegenheit gegenüber traditionellen Lehrformaten. Dies wurde zum Beispiel für den Erwerb von Wissen und die Anwendung von Fähigkeiten zu Erster Hilfe (Kononowicz et al. 2012), zu Atemwegserkrankungen bei pädiatrischen Patienten (LeFlore et al. 2012), zur Abschätzung der Behandlungsdringlichkeit von Unfallopfern (Knight et al. 2010) sowie zu kardiologischen Erkrankungen einschließlich Inspektion und Auskultation kardiologischer Patienten (Vukanovic-Criley et al. 2008) gezeigt.

Die Akzeptanz und Motivation von Studierenden zur Nutzung von Computersimulationen ist derzeit hoch, da sie ein höheres Interesse an den Lehrinhalten wecken als traditionelle Lehrformate (Kaczmarczyk et al. 2015; Dankbaar et al. 2016a; Dankbaar et al. 2017) und als wertvolle Hilfe zur Wissenskonsolidierung eingeschätzt werden (Bryce et al. 1998; Kaczmarczyk et al. 2015).

1.5 Prüfungen

Leistungsnachweise sind zentraler Bestandteil der Lehre und verfolgen sowohl persönliche als auch institutionelle und gesellschaftliche Ziele (Epstein und Hundert 2002; Cilliers et al. 2010). Sie dienen nicht nur der Bewertung des Leistungsstands von Studierenden, sondern sind zusätzlich wichtig unter anderem für die Bedeutung der Lehrinhalte aus Sicht der Studierenden („*Assessment drives learning*“ (McLachlan 2006)) und für die Bescheinigung ärztlicher Kompetenz von Absolventen aus Sicht der Öffentlichkeit (Epstein und Hundert 2002).

1.5.1 Prüfungskonsequenzen

In der Lehre werden im Allgemeinen summative und formative Prüfungen eingesetzt (Anziani et al. 2008). Summative Prüfungen („*Assessment of learning*“) dienen der (benoteten) Bewertung von Leistung und Lernerfolg (Roediger und Karpicke 2006). Sie werden gewöhnlich am Ende einer Lehreinheit durchgeführt, um das Erreichen angestrebter Lernziele zu überprüfen und Studierenden den Übergang in einen nächsten Abschnitt des Curriculums zu gewähren bzw. zu verweigern (Anziani et al. 2008). Im Gegensatz dazu werden formative Prüfungen („*Assessment for learning*“ (Roediger und Karpicke 2006)) regelmäßig während einer Lehreinheit eingesetzt und ermöglichen eine (unbenotete)

Rückmeldung über Lernstand und –fortschritt (Hudson 2006; Anziani et al. 2008). Auf Basis der Rückmeldung können Studierende Probleme im Lernprozess identifizieren (Anziani et al. 2008) und Lernstrategien anpassen, um ihren Lernprozess zu optimieren und die angestrebten Lernziele im Laufe der Lehreinheit zu erreichen (Hudson 2006).

1.5.2 Prüfungsformate

In der ÄApprO finden sich Bestimmungen bezüglich der Meldung und Zulassung zu Prüfungen sowie der Art, Bewertung und Durchführung von Prüfungen (Güntert 2003). Demnach müssen im vorklinischen Studienabschnitt 15 Leistungsnachweise und zusätzlich ein benoteter Leistungsnachweis in einem Wahlfach erbracht werden. Im klinischen Studienabschnitt müssen benotete Leistungsnachweise in 21 Hauptfächern, 14 Querschnittsbereichen, einem Wahlfach und fünf Blockpraktika erbracht werden. Die Benotung wird im „Zeugnis der Ärztlichen Prüfung“ gesondert zu den Prüfungsergebnissen aus dem Ersten und Zweiten Staatsexamen ausgewiesen (Güntert 2003).

Die Bestimmungen der ÄApprO werden für die Medizinische Fakultät Göttingen in ihrer Studienordnung konkretisiert (Studienordnung für den Studiengang Humanmedizin an der Georg-August-Universität Göttingen). Nach ihren Richtlinien dürfen Erfolgskontrollen mündlich (z. B. als Referat, mündliche Prüfung), schriftlich (z. B. als ‚*Multiple-Choice*‘-Prüfung, ‚*Key-Feature*‘-Prüfung) und praktisch (z. B. als objektiv strukturierte klinische Überprüfung (‚*OSCE*‘)) durchgeführt werden.

Zur Prüfung von theoretischem Faktenwissen (d. h. deskriptives Wissen) sind etablierte Prüfungsformate wie z. B. ‚*Multiple-Choice*‘-Prüfungen geeignet (van der Vleuten und Driessen 2014). Im häufig verwendeten „Fragentyp A“ muss die einzig richtige bzw. falsche Antwort unter fünf Alternativen ausgewählt werden, wodurch das Erinnern bzw. Wiedererkennen von Faktenwissen geprüft wird (Shumway und Harden 2003). Andere Fragentypen verlangen die voneinander unabhängige Bewertung (richtig/falsch) mehrerer Aussagen oder die korrekte Zuordnung von Feststellungen und ergänzenden Aussagen (Shumway und Harden 2003). Unabhängig vom Fragentyp geht die Prüfung von Kompetenzen allerdings selten über die Taxonomiestufen „Kenntnis“ und „Verständnis“ nach Bloom (1984) hinaus (Biggs 1996). Diese oberflächliche Art des Prüfens kann die Lernstrategien von Studierenden dahingehend beeinflussen, dass Faktenwissen lediglich auswendig gelernt (Van der Vleuten und Newble 1995) und ‚tiefes Lernen‘ verhindert wird (Shumway und Harden 2003; McLachlan 2006).

Soll durch Prüfungen allerdings Handlungs- und Begründungswissen (d. h. prozedurales Wissen), zum Beispiel in Form von ‚*clinical reasoning*‘, erfasst werden, so müssen Prüfungsformate gewählt werden, welche die Anwendung von Wissen und Problemlösefähigkeiten im klinischen Kontext erfordern (Biggs 1996). Hier bieten sich unter anderem ‚*Key Feature*‘-Prüfungen an, die eine hohe Konstruktvalidität im Hinblick auf die Überprüfung entsprechender Kompetenzen aufweisen (Hrynchak et al. 2014). Ein ‚*Key Feature*‘ ist definiert als „*critical step in the successful resolution of a problem*“ (Page und Bordage 1995). ‚*Key Feature*‘-Prüfungen konzentrieren sich auf die Abfrage dieser kritischen Schritte und Schlüsselentscheidungen zum Lösen eines klinischen Problems (Page und Bordage 1995). ‚*Key Feature*‘-Fälle beginnen mit einer kurzen Darstellung des klinischen Szenarios (u. a. Patientendaten und Konsultationsgrund). Es folgt eine variable Anzahl an Fragen zur Prüfung der ‚*Key Features*‘, welche sich typischerweise auf die Interpretation von Symptomen, auf Differentialdiagnosen, auf Untersuchungen, auf das Management oder auf Therapiemöglichkeiten beziehen (Farmer und Page 2005). Die Beantwortung der Fragen erfolgt entweder im „*Write In*“-Format, bei dem die richtige Antwort in ein freies Textfeld eingetragen werden muss, im „*Short Menu*“-Format, bei dem die Antwort aus einer Liste mit etwa 15-45 vorgegebenen Alternativen ausgesucht werden muss oder im „*Long Menu*“-Format, welches auf einer Liste mit hunderten Antwortmöglichkeiten basiert (Page und Bordage 1995). Letzteres kommt einem offenen Fragenformat nahe, da aufgrund der hohen Anzahl an Alternativen die Antwort eigenständig generiert und anschließend in der Antwortliste nachgeschlagen werden muss (Kopp et al. 2006).

Nicht zuletzt werden im Medizinstudium auch praktische und kommunikative Fertigkeiten auf der Ebene von Handlungskompetenz gelehrt. Zu ihrer Überprüfung bedarf es Prüfungsformaten, in denen die klinisch-praktische Anwendung von Fertigkeiten demonstriert werden muss. Hierzu zählen beispielsweise objektiv strukturierte klinische Überprüfungen (*OSCE*), in denen an mehreren Stationen klinisch-praktische Fertigkeiten (z. B. Untersuchungstechniken) und kommunikative Fähigkeiten (z. B. Anamneseerhebung) anhand einer Checkliste beurteilt werden (Shumway und Harden 2003; Nikendei und Jünger 2006).

1.6 Ziele dieser Arbeit

‚*Clinical reasoning*‘ stellt eine zentrale und zur selbstständigen Tätigkeit unerlässliche ärztliche Kompetenz dar, deren Vermittlung Bestandteil der medizinischen Lehre sein sollte. Klassischerweise erfolgt der Unterricht (auch derart komplexer Lernziele) in Form von

Vorlesungen, Seminaren oder Kleingruppenunterricht. Ziel dieser Arbeit war es, das ‚*Serious Game*‘ EMERGE, eine virtuelle Notaufnahme, als innovative Methode zur Vermittlung ärztlicher Kompetenzen und zum Training von ‚*clinical reasoning*‘-Fähigkeiten in der curricularen Lehre zu implementieren.

Es wurde beabsichtigt, mithilfe der hier vorgestellten Studie folgende Studienfragen zu beantworten:

- 1) Erzielen Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, bessere bzw. schlechtere Ergebnisse im Hinblick auf differentialdiagnostische und –therapeutische Kompetenzen als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts lernen?
- 2) Inwieweit sind das im Rahmen eines problemorientierten Kleingruppenunterrichts oder EMERGE erlernte diagnostische und therapeutische Wissen sowie die Kompetenzen hinsichtlich Zeitmanagement und Priorisierung fallspezifisch und/oder formatspezifisch?
- 3) Gehen Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, im Behandlungsablauf von virtuellen Patienten strukturell anders vor als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts diskutieren?
- 4) Empfinden Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, bei der gleichzeitigen Behandlung von mehreren virtuellen Patienten weniger Stress als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts diskutieren?
- 5) Halten Studierende die Computer-Simulation EMERGE für ein geeignetes Lernformat zum Erwerb ärztlicher Kompetenzen?

2 Methoden

2.1 Studiendesign

Bei der hier vorgestellten Studie handelt es sich um eine prospektive, nicht randomisierte Interventionsstudie, an der im Sommersemester 2016 Studierende der Humanmedizin im sechsten klinischen Semester an der UMG teilnehmen konnten.

Im folgenden Kapitel wird zunächst die Organisation und Durchführung der Studie beschrieben. Anschließend werden die verwendeten Lehrformate, die Erhebungsinstrumente und die Datensammlung sowie die statistischen und qualitativen Auswertungsmethoden erläutert. Das Studienprotokoll dieses Forschungsprojekts wurde bei der Ethikkommission mit der Antragsnummer 18/3/16 eingereicht und genehmigt (positives Votum vom 22.04.2016).

2.2 Organisation und Durchführung

Unabhängig von der hier vorgestellten Studie fand im Sommersemester 2016 von der ersten bis zur sechsten Semesterwoche (11.04.2016 – 20.05.2016) erstmalig das Repetitionsmodul 6.x für das sechste klinische Semester an der UMG statt. Im Rahmen dieses sechswöchigen Moduls nahmen die Studierenden entweder an EMERGE (ein ‚*Serious Game*‘) oder an POLEMA („Problem orientiertes Lernen – elektronische Medien assistiert“) teil. Hierzu wurden die Studierenden einige Wochen vor Beginn des sechsten klinischen Semesters über das neue Repetitionsmodul 6.x und die zur Auswahl stehenden Lehrformen EMERGE und POLEMA informiert. Die Teilnahme an einer der beiden Lehrformen war für alle Studierenden verpflichtend, wobei der anwesenheitspflichtige Zeitaufwand für beide Formate gleich war (maximal 20% Fehltermine). Auf Grundlage von Wunsch und zur Verfügung stehenden Kapazitäten wurden die Studierenden anschließend einer der beiden Lehrformen zugeteilt. Da beide Lehrformate in das didaktische Gesamtkonzept des neuen Moduls eingebettet waren, wurden thematisierte Inhalte nicht nur dort, sondern auch in Vorlesungen und Seminaren angesprochen („*blended learning*“ s. o.).

Das Repetitionsmodul 6.x verfolgte das übergeordnete Lernziel, ‚*clinical reasoning*‘-Fähigkeiten zu trainieren. Die Studierenden sollten die Möglichkeit erhalten, ihre differentialdiagnostischen Kompetenzen bei klinisch relevanten Symptomkomplexen zu steigern. Dies bedeutet zunächst, die relevanten Differentialdiagnosen ausgehend von einem bestimmten Symptom nennen zu können. Darüber hinaus gilt es aber auch, die Dringlichkeit klinischer Situationen einzuschätzen und die relevanten diagnostischen und therapeutischen

Schritte einzuleiten.

Am ersten Tag des Moduls fand eine Einführungsveranstaltung statt, um die Studierenden über den organisatorischen Ablauf und die Lehrinhalte des Repetitionsmoduls zu informieren. Im Zuge dieser Veranstaltung wurde allen Teilnehmern ein Modulbuch ausgehändigt, in welchem die zentralen Inhalte des Moduls identifiziert und in Form von Lernzielen festgehalten wurden (z. B. „Zu Modulende können Sie das diagnostische Vorgehen bei einem/r Patienten/in mit Bauchschmerzen fachgerecht planen.“). Das Modulkript wurde auf Grundlage des Göttinger Lernzielkatalogs (s. o.) entwickelt.

Am Ende des sechswöchigen Moduls wurde eine formative (d. h. nicht benotete) Prüfung wesentlicher Modulinhalte durchgeführt. Die Studierenden bearbeiteten sechs ‚*Key Feature*‘-Fälle (siehe Kapitel 2.4.3) und vier Fälle in EMERGE (siehe Kapitel 2.3). Zudem füllten die Studierenden einen Fragebogen aus.

Im Anschluss an das Modul 6.x hatten alle Teilnehmer die Möglichkeit, das ‚*Serious Game*‘ EMERGE an drei definierten Terminen in seinem vollen Umfang zu nutzen, unabhängig davon, ob sie zuvor an EMERGE oder an POLEMA teilgenommen hatten. Hierdurch wurde allen interessierten Studierenden die Möglichkeit gegeben, das neu eingeführte ‚*Serious Game*‘ zu nutzen, auch wenn sie sich im Modul 6.x zu einer Teilnahme an POLEMA entschieden hatten.

2.2.1 Probandenrekrutierung

Eine Woche vor Beginn des Sommersemesters 2016 wurden alle für das Modul 6.x angemeldeten Studierenden per Email über die geplante Studie informiert. Im Zuge dessen erhielten die Studierenden eine Probandeninformation, in der Hintergrund und Ablauf der Studie erklärt wurden, sowie eine Einverständniserklärung, mit deren Unterzeichnung sie der Auswertung ihrer Daten zustimmen konnten.

In der Einführungsveranstaltung am ersten Tag des Moduls wurde die Studie ein weiteres Mal vorgestellt und sowohl die Probandeninformation als auch die Einverständniserklärung in Papierform ausgeteilt. Die Studierenden wurden ausdrücklich über die Freiwilligkeit der Teilnahme an der Studie und die Möglichkeit zum Widerruf der Einwilligung ohne Angabe von Gründen aufgeklärt. Außerdem wurden die Studierenden eingeladen, Fragen zur Studie zu stellen.

2.3 Lehrformen

Im folgenden Abschnitt werden die beiden Lehrformate POLEMA und EMERGE näher erläutert. Es wird sowohl auf die thematisierten Inhalte als auch auf die genaue praktische Durchführung eingegangen.

2.3.1 Kleingruppenunterricht POLEMA

Im Sommersemester 2016 fanden im Rahmen des Repetitionsmoduls 6.x für jeden Teilnehmer von POLEMA 11 Kurstage zu je zwei Lehrveranstaltungsstunden (entsprechend eineinhalb Zeitstunden) statt. Während der ersten 30 Minuten der Veranstaltungsreihe wurden den Studierenden Konzept und Ablauf der Veranstaltung vorgestellt. Ebenso wurden in den letzten 30 Minuten der Veranstaltungsreihe keine Unterrichtsinhalte thematisiert, sondern eine Abschlussbesprechung und eine Evaluation durchgeführt.

Die Studierenden arbeiteten im Rahmen von POLEMA in festen Gruppen zu je 6-8 Studierenden zusammen. Zur Durchführung der Unterrichtseinheiten kamen jeweils drei bis vier Gruppen parallel in den Räumen des STÄPS (Studentisches Trainingszentrum Ärztlicher Praxis und Simulation) zusammen. In jeder Gruppe war ein fester studentischer Tutor für die Moderation der Kleingruppenarbeit und die Steuerung der Computerdatenbank, in welcher die Patientenfälle hinterlegt sind, verantwortlich. Alle eingesetzten studentischen Tutoren haben im Vorfeld an einer Schulung durch die POLEMA-Koordination teilgenommen. Im Rahmen dieser Veranstaltung wurden sie in die Organisation und Methodik von POLEMA eingeführt, auf die Moderation der Gruppenarbeit vorbereitet und im Umgang mit der Computerdatenbank geschult.

Im Rahmen von POLEMA wurden im Sommersemester 2016 insgesamt fünf Patientenfälle behandelt (siehe Tabelle 2). Für die Bearbeitung eines jeden Patientenfalls standen in etwa zweieinhalb Kurstage zur Verfügung. Während am jeweils ersten Kurstag der Patientenfall vorgestellt und wesentliche Differentialdiagnosen thematisiert wurden, sollten am zweiten Kurstag diagnostische und therapeutische Schritte zur Behandlung des Patienten erarbeitet werden. Die Zeit zwischen den Kursterminen wurde zum Selbststudium der thematisierten Inhalte genutzt. Im Anschluss kamen in der ersten Hälfte des jeweils dritten Kurstages alle Gruppen zur Besprechung der Ergebnisse im Plenum zusammen. Jedes Plenum wurde von einem ärztlichen Fallexperten moderiert, welcher den Studierenden Hintergrundinformationen lieferte und zur Diskussion bereitstand.

Tabelle 2: Inhalte von POLEMA im Sommersemester 2016

Kurstag 1	Einführung	Kurstag 7	Plenum Morbus Hodgkin
	Pankreas-Karzinom		Magen-Karzinom
Kurstag 2	Pankreas-Karzinom	Kurstag 8	Magen-Karzinom
Kurstag 3	Plenum Pankreas-Karzinom	Kurstag 9	Plenum Magen-Karzinom
	Sarkoidose		Morbus Wegener
Kurstag 4	Sarkoidose	Kurstag 10	Morbus Wegener
Kurstag 5	Plenum Sarkoidose	Kurstag 11	Plenum Morbus Wegener
	Morbus Hodgkin		Abschlussevaluation
Kurstag 6	Morbus Hodgkin		

2.3.2 Computersimulation EMERGE

EMERGE ist ein ‚*Serious Game*‘, das von der UMG in Zusammenarbeit mit der PatientZeroGames GmbH und dem Universitätsklinikum Hamburg-Eppendorf entwickelt wurde (siehe Bildschirmaufnahmen in Abbildung 1). Es handelt sich um eine virtuelle Notaufnahme mit acht normalen Bettenplätzen und zwei Schockräumen. In der 3D-Computersimulation nimmt der Spieler die Rolle des verantwortlichen Arztes ein und behandelt gleichzeitig bis zu zehn virtuelle Patienten. Dabei sind alle wesentlichen Aspekte der Tätigkeit in einer Notaufnahme abgebildet. Im Folgenden werden die wesentlichen Spielfunktionen zusammengefasst.

Trifft ein neuer Patient in der Notaufnahme ein, erscheint eine Meldung, in welcher Name, Geschlecht und Konsultationsanlass bzw. Symptomkomplex des neuen Patienten genannt werden (z. B. „Manuela Hertzog ist mit folgenden Beschwerden in die Notaufnahme gekommen: Bauchschmerzen und Fieber“). Der Spieler muss nun zunächst die relevanten Differentialdiagnosen eines Symptomkomplexes wie „Luftnot“, „Brustschmerz“ oder „Fieber“ kennen und die Dringlichkeit der klinischen Situation einschätzen. Zur weiteren Behandlung steht anschließend das Auswahlmenü mit sechs Unterpunkten (Anamnese, Untersuchung, Behandlung, Labor, Diagnose und Verlegung) zur Verfügung.

- Anamnese: Im Anamnesemenü sind mehr als 70 Fragen zu finden. Für jede Frage sind assoziierte Stichworte, Synonyme und Abkürzungen hinterlegt worden, durch deren Eingabe in eine Suchmaske die jeweilige Frage erscheint.
- Untersuchung: Während grundlegende Diagnostik, z. B. die körperliche Untersuchung, vom

Spieler in Person des Arztes selbst am Patientenbett durchgeführt wird, werden andere Untersuchungen (z. B. Blutzuckerbestimmung) vom Pflegepersonal ausgeführt und ermöglichen es dem Spieler, währenddessen an anderer Stelle weiterzuarbeiten. Auch aufwändige apparative Untersuchungen (z. B. eine Ösophagogastroduodenoskopie) sind möglich. Hierzu wird der Patient entweder für eine gewisse Zeit in die entsprechende Abteilung verlegt oder ein Spezialist führt die angeforderte Untersuchung auf der Notaufnahme durch. Außerdem kann ein ärztliches Konsil einer bestimmten Fachrichtung eingeholt werden, um einen kompetenten Rat zu erhalten. Alle Untersuchungsergebnisse werden in der Patientenakte dokumentiert, wo sie bis zum Fallabschluss jederzeit einsehbar bleiben.

- Labor: Der Spieler kann ca. 80 Laborparameter einzeln auswählen. Zudem können Laborprofile, Blutgasanalysen oder Urinuntersuchungen angefordert werden.
- Diagnose: Es findet sich eine alphabetische Auflistung von ca. 130 verschiedenen Diagnosen. Das Auswählen einer Diagnose ist spätestens zum Zeitpunkt des Fallabschlusses erforderlich; die Diagnose kann aber auch schon frühzeitig im Sinne einer Verdachtsdiagnose gestellt und später gegebenenfalls angepasst werden.
- Behandlung: Im Behandlungsmenü kann der Spieler zwischen apparativen Behandlungen (z. B. Anlage einer Thoraxdrainage) und der Gabe von über 150 Medikamenten wählen, welche in verschiedenen Darreichungsformen verabreicht werden können.
- Verlegung: Der Patient kann schließlich sowohl auf verschiedene Stationen (Normal-, ‚Intermediate-Care‘ oder Intensivstation) als auch in den OP verlegt oder aus dem Krankenhaus entlassen werden.

Im Anschluss an die Verlegung kann der Spieler eine Rückmeldung einsehen, in welcher die von ihm unternommenen Tätigkeiten den wichtigsten tatsächlich notwendigen Tätigkeiten gegenübergestellt werden. Außerdem wird in Form eines Oberarztcommentars eine korrekte Vorgehensweise bezüglich des vorgestellten Krankheitsbildes vorgeschlagen.

Durch EMERGE werden nicht nur inhaltliche Kenntnisse, sondern auch organisatorische Fähigkeiten geschult, da zur Gewährleistung einer effizienten Patientenversorgung bei der parallelen Patientenbehandlung Prioritäten gesetzt und Dringlichkeiten richtig eingeschätzt werden müssen. Die Krankheitsbilder und die Anzahl der eintreffenden Patienten können im Vorfeld an den Ausbildungsstand angepasst werden, sodass EMERGE Studierenden in jedem klinischen Studienabschnitt die Möglichkeit bietet, ärztliche Kompetenzen zu trainieren.



Abbildung 1: Bildschirmaufnahmen von EMERGE

Oben links: Benachrichtigung über die Ankunft eines neuen Patienten in der virtuellen Notaufnahme. Oben rechts: Patientenübergabe durch Rettungsdienstmitarbeiter. Unten links: Auswahlmennü für Laboranforderungen. Unten rechts: Auswahlmennü für Anamnesefragen.

2.3.2.1 Inhalt von EMERGE

Für die Integration der virtuellen Notaufnahme EMERGE in die modulare Lehre der UMG wurden im Vorfeld 50 verschiedene Patientenfälle konzipiert. Diese repräsentieren wesentliche Gesundheitsstörungen des Göttinger Lernzielkatalogs (s. o.) und stammen aus den internistischen Fachbereichen Kardiologie/Pulmologie, Nephrologie/Rheumatologie, Gastroendokrinologie und Hämatologie/Onkologie (s. Tabelle 3).

Tabelle 3: EMERGE-Fälle, die im Rahmen der Studie eingesetzt wurden.

EMERGE-FÄLLE			
Kardiologie/Pulmologie	Nephrologie/Rheumatologie	Gastroendokrinologie	Hämatologie/Onkologie
Synkope	Urolithiasis	Alkoholintoxikation	obere gastrointestinale Blutung
STEMI	Pyelonephritis	Hypoglykämie	Armvenerthrombose
Pneumothorax	Gichtanfall	Reflux	EBV Infektion
Aortendissektion	Sarkoidose	Bronchialkarzinom	Morbus Hodgkin
Herzinsuffizienz	Hypertonus	Leberzirrhose	multiples Myelom
Vorhofflimmern	Harnwegsinfekt	Hyperthyreose	Fieber in Aplasie
Bronchitis	Urosepsis	Cholezystolithiasis	Thrombopenie
COPD	akutes Nierenversagen	Gastroenteritis	Leukämie
Lungenarterienembolie, stabil	chronisches Nierenversagen	Divertikulitis	Eisenmangelanämie
Lungenarterienembolie, instabil	Hyponatriämie	Schwangerschaft	Tumorlysesyndrom
AVNRT	nephrotisches Syndrom	Medikamentenintoxikation	Hyperkalziämie
Pneumonie		Ketoazidose	
Asthmaanfall		Pankreatitis	
NSTEMI		untere gastrointestinale Blutung	

2.3.2.2 Einsatz von EMERGE im Modul 6.x

Die 3D-Computersimulation EMERGE wurde im Repetitionsmodul 6.x im Sommersemester 2016 erstmalig eingesetzt. Für jeden Teilnehmer fanden über das sechswöchige Modul verteilt zehn Termine zu je 90 Minuten statt. Der elfte Termin entsprach einer 90-minütigen formativen Prüfung. An diesem Termin bestand für alle Studierenden des Moduls 6.x Anwesenheitspflicht, unabhängig von ihrer Teilnahme an EMERGE oder POLEMA.

Alle EMERGE-Termine fanden im Digitalen Prüfungs- und Schulungszentrum (DiPS) der UMG statt, welches aus drei Räumen mit insgesamt über 200 Computerplätzen besteht. Für EMERGE stand lediglich das DiPS 3 (Raum 257) zur Verfügung, welches 49 Computerplätze umfasst, die zur Darstellung der 3D-Bilddaten von EMERGE vor Beginn des Sommersemesters 2016 mit neuen, 3D-fähigen ‚Clients‘ ausgestattet wurden. Die Aufbereitung der Bilddaten erfolgte auf Servern mit beschleunigten Grafikkarten, die ebenso wie die ‚Clients‘ eigens für den Einsatz von EMERGE angeschafft wurden.

Aufgrund der Raumkapazitäten musste jeder EMERGE-Termin drei Mal angeboten werden. Hierzu wurden drei konstante und etwa gleich große Studierendengruppen (die EMERGE-Gruppen A, B und AB) gebildet. Allen drei Gruppen wurden virtuelle Patienten mit ähnlichen Konsultationsanlässen und Symptomen (z. B. „Fieber“, „Bewusstseinsstörung“)

präsentiert, bei denen jeweils verschiedene Differentialdiagnosen in Frage kommen. Die tatsächlich zugrundeliegenden Erkrankungen stammten in Gruppe A aus den Fachbereichen Kardiologie/Pulmologie und Nephrologie/Rheumatologie und in Gruppe B aus den Fachbereichen Gastroendokrinologie und Hämatologie/Onkologie. In der Gruppe AB stammten die zugrundeliegenden Erkrankungen aus allen vier Fachbereichen.

2.3.2.3 Gestaltung der einzelnen EMERGE-Termine

Zur Gestaltung der einzelnen EMERGE-Termine wurde im Vorfeld ein detaillierter Plan erarbeitet. Hierzu wurden die Patientenfälle eines jeden Fachbereichs zunächst in zwei Fallgruppen zu je fünf bis sechs Fällen aufgeteilt (siehe Tabelle 4). Davon ausgeschlossen waren die vier folgenden Patientenfälle: Der „Asthmaanfall“ (Kardiologie/Pulmologie), der „NSTEMI“ (Kardiologie/Pulmologie), die „Pankreatitis“ (Gastroendokrinologie) und die „Untere gastrointestinale Blutung“ (Gastroendokrinologie). Diese Fälle wurden erstmalig in der Prüfung präsentiert, die in der Mitte der sechsten Modulwoche stattfand.

Tabelle 4: Einteilung der EMERGE-Fälle in acht Fallgruppen

EMERGE-FALLGRUPPEN					
EMERGE-Gruppe A	EMERGE-Gruppe AB	Kardiologie/Pulmologie		Nephrologie/Rheumatologie	
		Fallgruppe A1	Fallgruppe A2	Fallgruppe A3	Fallgruppe A4
		Synkope	Bronchitis	Urolithiasis	Harnwegsinfekt
		STEMI	COPD	Pyelonephritis	Urosepsis
		Pneumothorax	Lungenarterienembolie, stabil	Gichtanfall	Akutes Nierenversagen
		Aortendissektion	Lungenarterienembolie, instabil	Sarkoidose	Chronisches Nierenversagen
		Herzinsuffizienz	AVNRT	Hypertonus	Hyponatriämie
		Vorhofflimmern	Pneumonie		Nephrotisches Syndrom
EMERGE-Gruppe B	EMERGE-Gruppe AB	Gastroendokrinologie		Hämatologie/Onkologie	
		Fallgruppe B1	Fallgruppe B2	Fallgruppe B3	Fallgruppe B4
		Alkoholintoxikation	Cholezystolithiasis	Obere gastrointestinale Blutung	Fieber in Aplasie
		Hypoglykämie	Gastroenteritis	Armvenenthrombose	Thrombopenie
		Reflux	Divertikulitis	EBV Infektion	Leukämie
		Bronchialkarzinom	Schwangerschaft	Morbus Hodgkin	Eisenmangelanämie
		Leberzirrhose	Medikamentenintoxikation	multiples Myelom	Tumorlysesyndrom
		Hyperthyreose	Ketoazidose		Hyperkalziämie
PRÜFUNG					
EMERGE A/B/AB	Asthmaanfall	NSTEMI	Untere gastrointestinale Blutung	Pankreatitis	

Anschließend wurde für jeden der insgesamt 10 Termine festgelegt, welche Fallgruppen der

jeweiligen Gruppe präsentiert werden (siehe Tabelle 5). Hierzu wurde für jeden Termin ein bestimmtes Spielszenario entworfen, in welchem jeweils 1-8 Fallgruppen berücksichtigt worden sind. Die Szenarien wurden vor jedem Termin in den EMERGE-Einstellungen, genauer in der Optionsdatei, hinterlegt.

Das Programm EMERGE wurde an jedem Termin vom Dozentenplatz aus auf allen Computern gestartet. Hierzu wurde die Klassenraumsteuerung „NetMan for Schools“ (H+H Software GmbH, Göttingen) genutzt, die das zentrale Starten und Beenden von Anwendungen auf einzelnen bzw. allen Computern des DiPS ermöglicht und den Internetzugriff unterbinden kann. Nach Eingabe der Matrikelnummer durch den Studierenden wurde beim Spielstart die Nutzung des hinterlegten Szenarios forciert, sodass nur die im Voraus ausgewählten Fallgruppen auf der virtuellen Notaufnahme eintrafen. Die Abfolge des Eintreffens der jeweiligen Patientenfälle einer Fallgruppe erfolgte zufällig. Jede für eine Gruppe ausgewählte Fallgruppe wurde im Verlauf der ersten vier Modulwochen mindestens einmal wiederholt. In der fünften Woche konnten Fälle aller für die jeweilige Gruppe ausgewählten Fallgruppen auf der Notaufnahme eintreffen.

Tabelle 5: Verteilung der EMERGE-Fallgruppen auf die einzelnen Termine der drei EMERGE-Gruppen

EMERGE-FALLGRUPPEN: VERTEILUNG																	
		Gruppe A				Gruppe B				Gruppe AB							
Woche 1	Termin 1	A1				B1				A1	B1						
	Termin 2	A1				B1				A1	B1						
Woche 2	Termin 3	A2				B2				A2	B2						
	Termin 4	A2				B2				A2	B2						
Woche 3	Termin 5	A3	A1			B3	B1			A3	B3	A1					
	Termin 6	A3	A2			B3	B2			A3	B3	B1					
Woche 4	Termin 7	A4	A3			B4	B3			A4	B4	A3					
	Termin 8	A4	A3			B4	B3			A4	B4	B3					
Woche 5	Termin 9	A1	A2	A3	A4	B1	B2	B3	B4	A1	A2	A3	A4	B1	B2	B3	B4
	Termin 10	A1	A2	A3	A4	B1	B2	B3	B4	A1	A2	A3	A4	B1	B2	B3	B4
Woche 6	PRÜFUNG																

Alle EMERGE-Termine wurden von einem ärztlichen Mitarbeiter der UMG geleitet, der in

die Entwicklung des Spiels und der Fälle involviert war. Dieser führte die Studierenden an ihrem jeweils ersten Termin in die Steuerung des Spiels ein und leitete sie anschließend an, EMERGE weitestgehend selbstständig zu spielen. Darüber hinaus wurden die Studierenden dazu angeregt, die Inhalte der präsentierten Patientenfälle im Selbststudium vor- und nachzubereiten. Hierzu wurden für alle Patientenfälle aktuelle Behandlungsleitlinien der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. als PDF-Dokumente bereitgestellt, auf welche die Studierenden während der EMERGE-Termine jederzeit zugreifen konnten. Zudem wurde den Studierenden geraten, die präsentierten Inhalte in einschlägigen Lehrbüchern nachzulesen. Um Fragen auch im persönlichen Gespräch klären zu können, stand ihnen der ärztliche Betreuer während aller Termine zum Gespräch zur Verfügung. Weiterhin wurden die Studierenden eingeladen, am Ende eines jeden EMERGE-Termins das im Modulbuch enthaltene Lerntagebuch auszufüllen (siehe Anhang 1). Diese freiwillige Maßnahme sollte die Studierenden in ihrem konstruktivistischen Lernprozess unterstützen (Biggs 1996), indem sie Gelerntes selbstständig zusammenfassen und weiterführende Fragen formulieren.

2.3.2.4 Log-Dateien

Das ‚*Serious Game*‘ EMERGE speichert automatisch alle Aktivitäten eines jeden Spielers in Log-Dateien. Hierbei handelt es sich pro Spieltermin und Spieler um jeweils drei Dateien: Erstens eine TXT-Datei (reine Textdatei); zweitens eine CSV-Datei (*Comma-Separated Values*), die beispielsweise mit dem Programm Microsoft Excel geöffnet werden kann; drittens eine XML-Datei (*Extensible Markup Language*), welche zum Beispiel mithilfe eines Webbrowsers geöffnet werden kann. Die Log-Dateien wurden nach jedem EMERGE-Termin auf dem Server abgelegt und anschließend von einem Mitarbeiter der Lehr-IT (Informationstechnologie) des Bereichs Medizindidaktik und Ausbildungsforschung an den Studienleiter übergeben. Es wurden nur Log-Dateien von Studierenden ausgewertet, die zuvor anhand der Einwilligungserklärung der Auswertung ihrer Daten zugestimmt hatten.

2.4 Die formative Abschlussprüfung

2.4.1 Organisation der Abschlussprüfung

In der letzten Modulwoche wurde unabhängig von der Studie eine obligate 90-minütige formative Abschlussprüfung durchgeführt, die unter anderem den Koordinatoren des neu etablierten Moduls 6.x dazu diente, den studentischen Lernerfolg hinsichtlich des differentialdiagnostischen und –therapeutischen Vorgehens abzuschätzen. Die

studentischen Leistungen in dieser Prüfung gingen in die Analysen der vorliegenden Promotionsarbeit ein, sofern die Studierenden einer entsprechenden Datennutzung zugestimmt hatten. Im ersten Teil der Prüfung mussten die Studierenden in einem Zeitraum von 30 Minuten sechs ‚*Key Feature*‘-Fälle lösen, im zweiten Teil mussten innerhalb von 45 Minuten vier Fälle im EMERGE-Format bearbeitet werden. Außerdem bearbeiteten die Studierenden einen Fragebogen, in dem unter anderem die Nutzung digitaler Medien thematisiert und um eine inhaltliche Bewertung der virtuellen Notaufnahme gebeten wurde.

2.4.2 Fragebögen

Die verwendeten Fragebögen wurden speziell für dieses Forschungsvorhaben entwickelt und mithilfe von EvaSys® (Electric Paper Evaluationssysteme GmbH, Lüneburg) online erstellt. Die Beantwortung der Fragebögen erfolgte anonym und papierbasiert; hierzu wurde zu Beginn der Prüfung ein Exemplar an jeden Studierenden ausgeteilt.

Die Fragebögen gliederten sich in vier Teilbereiche und umfassten insgesamt 26 Fragen (siehe Anhang 2). Im ersten Abschnitt wurden soziodemografische Fragen zu Geschlecht, Alter und zur Vorerfahrung im Rettungswesen gestellt. Im zweiten Abschnitt ging es um die Nutzung digitaler Medien, die Vorerfahrung der Studierenden mit Lernspielen für das Medizinstudium und ihr subjektiv empfundenes Stresslevel während der EMERGE-Prüfungssitzung. In beiden Abschnitten wurden vorwiegend geschlossene Fragen gestellt. Das bedeutet, dass der Studierende unter zwei oder mehr Antwortalternativen die für ihn zutreffendste Antwort auswählen konnte. Der dritte Teil des Fragebogens konnte nur von den EMERGE-Teilnehmern beantwortet werden, denn sie wurden gebeten, acht Aussagen zur Motivation ihrer Teilnahme an EMERGE anhand einer sechsstufigen Likert-Skala zu bewerten (1 = „trifft voll zu“, 6 = „trifft überhaupt nicht zu“). Auf diese Weise bewerteten die EMERGE-Teilnehmer beispielsweise folgende Aussagen: „Ich habe mich zur Teilnahme an EMERGE entschieden, ... weil ich mich für den Einsatz von Lernspielen im Medizinstudium interessiere“ oder „... weil ich dadurch einen geringeren Arbeitsaufwand als bei POLEMA erwartet habe“. Auch die abschließende inhaltliche Bewertung der virtuellen Notaufnahme erfolgte nur durch die EMERGE-Teilnehmer. Hier stand erneut eine sechsstufige Likert-Skala (1 = „trifft voll zu“, 6 = „trifft überhaupt nicht zu“) zur Bewertung verschiedener Aussagen zur Verfügung. Unter anderem bewerteten die Studierenden die Aussagen „Ich halte EMERGE zur Vermittlung ärztlicher Kompetenzen im Medizinstudium für geeignet“ oder „Durch das Spielen von EMERGE habe ich gelernt, die Wertigkeit verschiedener diagnostischer Schritte einschätzen zu können“.

Die Abgabe der ausgefüllten Fragebögen durch die Studierenden erfolgte am Ende der Prüfungssitzung und wurde als Ausdruck des Einverständnisses mit der Analyse der anonymen Daten gewertet. Die Fragebögen wurden anschließend eingescannt. Nach Analyse und Verarbeitung der Daten konnten diese als CSV-Datei (*Comma-Separated Values*) und als SAV-Datei heruntergeladen und mit dem Programm Microsoft Excel bzw. mit dem Statistikprogramm SPSS geöffnet werden.

2.4.3 *Key Feature*-Prüfung

Im ersten Teil der Abschlussprüfung bearbeiteten die Studierenden innerhalb von 30 Minuten sechs *Key Feature*-Fälle.

2.4.3.1 Konstruktion der *Key-Feature* Fälle

Die sechs *Key Feature*-Fälle mit jeweils vier Fragen („Items“) wurden im Vorfeld für diesen Einsatz vom Studienleiter konstruiert. Jeder *Key Feature*-Fall begann mit einer klinischen Fallvignette. In dieser wurden neben Patienteninformationen, wie Alter und BMI, auch der Konsultationsgrund sowie der Medikamentenstatus angegeben, sofern diese Informationen zur Bearbeitung des Falls relevant waren. Anschließend folgten jeweils vier *Key Feature*-Fragen, wobei den Studierenden nach der Beantwortung jeder Frage die richtige Lösung angezeigt wurde. Auf diese Weise wurde bei aufeinander aufbauenden Fragen auch nach Auswahl einer falschen Antwort die weitere Bearbeitung des *Key Feature*-Falls ermöglicht. Als Antwortformat wurde das *Long Menu*-Format genutzt.

Jeweils zwei der *Key Feature*-Fälle bezogen sich auf den inhaltlichen Schwerpunkt der EMERGE-Gruppe A (Kardiologie/Pulmologie und Nephrologie/Rheumatologie) bzw. den inhaltlichen Schwerpunkt der EMERGE-Gruppe B (Gastroendokrinologie und Hämatologie/Onkologie). In EMERGE-Gruppe AB (alle Fälle) waren demnach alle diese vier Fälle in den Wochen zuvor thematisiert worden. Im Gegensatz dazu waren die Inhalte aller vier Fälle in POLEMA zuvor kein Thema. Zwei weitere Fälle stammten aus der inhaltlichen Schnittmenge zwischen EMERGE und POLEMA. In Tabelle 6 ist zusammengefasst, welche Gruppen gegenüber dem Inhalt eines *Key Feature*-Falls während der Teilnahme an EMERGE bzw. POLEMA exponiert worden waren.

Tabelle 6: Exposition der Gruppen gegenüber den Inhalten der *Key Feature*-Fälle

<i>Key Feature</i> -Fall	Inhalt wurde während des Moduls 6.x thematisiert in:			
	EMERGE-Gruppe A	EMERGE-Gruppe B	EMERGE-Gruppe AB	POLEMA
1. Herzinsuffizienz	X		X	

2. Hyponatriämie	X		X	
3. Hyperthyreose		X	X	
4. Fieber in Aplasie		X	X	
5. Sarkoidose	X		X	X
6. Morbus Hodgkin		X	X	X

2.4.3.2 Präsentation der ‚Key Feature‘-Fälle

Nach der Konstruktion wurden die ‚Key Feature‘-Fälle in das „ItemManagementSystem“ (IMS, Umbrella Consortium for Assessment Networks, Heidelberg, Deutschland) eingegeben. Die auf diese Weise zusammengestellte ‚Key Feature‘-Prüfung aus insgesamt 24 Fragen wurde anschließend als QTI-Datei (*Question and Test Interoperability*), einem Format zum Austausch von Prüfungsfragen, extrahiert und in das CAMPUS-Prüfungssystem (Medizinische Fakultät der Universität Heidelberg) importiert. Die importierte Prüfung wurde anschließend mithilfe der „Management-Console“ des CAMPUS-Prüfungssystems auf die sogenannten ‚Clients‘, die 49 Computerplätze im DiPS 3, verteilt. Die Anmeldung erfolgte über den studentischen Nutzernamen und das jeweilige Passwort. Über den LDAP-Server (*Lightweight Directory Access Protocol-Server*) wurden Benutzername und Passwort miteinander abgeglichen. Anschließend konnten die Studierenden nach Eingabe einer zuvor generierten spezifischen PIN (Persönliche Identifikationsnummer) zeitgleich auf die ‚Key Feature‘-Prüfung zugreifen. Nach Ablauf der zur Verfügung stehenden Zeit von 30 Minuten wurde die Prüfung über die „Management-Console“ vom Dozentenplatz aus für alle Teilnehmenden zeitgleich beendet. Die Prüfungsdaten wurden als Excel-Datei exportiert, in der die Möglichkeit bestand, korrekte Lösungen nachträglich anzuerkennen, welche bei der Konzeption der jeweiligen Frage noch nicht bedacht worden waren.

2.4.4 EMERGE-Prüfung

Im zweiten Teil der Abschlussprüfung mussten die Studierenden vier Fälle im EMERGE-Format bearbeiten (siehe Tabelle 4). Diese vier Fälle waren nicht nur den POLEMA-Teilnehmern, sondern auch allen drei EMERGE-Gruppen unbekannt.

Die Studierenden bearbeiteten die EMERGE-Fälle an denselben Computerplätzen, die sie zuvor zur Bearbeitung der ‚Key Feature‘-Fälle genutzt hatten. Zum technischen Wechsel zwischen dem ‚Key Feature‘-Prüfungsformat und der EMERGE-Prüfung fand zwischen beiden Teilen eine 5-minütige Pause statt. Über den Dozentenplatz wurde anschließend mithilfe der Klassenraumsteuerung „NetMan for Schools“ (H+H Software GmbH, Göttingen) das Programm EMERGE auf allen Computerplätzen gleichzeitig gestartet. Anschließend

konnten sich die Studierenden mithilfe ihrer Matrikelnummer einloggen und mit dem Spiel beginnen. Die Teilnehmer der Gruppe POLEMA wurden vom Studienleiter kurz in die Steuerung des Spiels eingeführt. Hierzu wurde beispielhaft ein Patient auf die Notaufnahme aufgenommen und behandelt, was mithilfe der Übertragung auf eine Leinwand von den Studierenden mitverfolgt werden konnte.

Nachdem die Studierenden das Spiel gestartet hatten, wurde ein für alle Teilnehmer identisches Prüfungsszenario geladen. Sowohl die Reihenfolge der vier Fälle als auch der zeitliche Abstand zwischen dem Eintreffen der einzelnen Patienten auf der Notaufnahme war für alle Studierenden gleich. Während der Bearbeitung der EMERGE-Fälle wurden vom anwesenden Studienleiter keine inhaltlichen Fragen beantwortet. Die Studierenden durften allerdings Fragen zur Steuerung des Spiels stellen, um die POLEMA-Teilnehmer, die das Spiel zuvor nicht kannten, in dieser Hinsicht nicht zu benachteiligen. Nach Ablauf der Prüfungszeit von 45 Minuten wurde das Spiel vom Dozentenplatz aus für alle Studierenden zeitgleich beendet. Wie in allen vorherigen EMERGE-Terminen wurden die Spielstände jederzeit in den Log-Dateien protokolliert.

2.4.4.1 Auswertung der EMERGE-Log-Dateien

Insgesamt wurden 893 Log-Dateien (EMERGE-Teilnehmer) aus den 10 EMERGE-Seminarterminen und 112 Log-Dateien (EMERGE und POLEMA-Teilnehmer) aus der EMERGE-Abschlussprüfung ausgewertet. Zu diesem Zweck wurde ein spezifisches Kodierungssystem entwickelt. Dafür wurden für jeden der 50 EMERGE-Fälle die in Tabelle 7 dargestellten Variablen festgelegt. Alle Log-Dateien wurden mit dem Programm Microsoft Excel 2016 für Windows, Version 16.0 geöffnet. Anschließend wurden die darin enthaltenen Informationen in das Statistikprogramm IBM SPSS Statistics für Windows, Version 24 (IBM Corp., Armonk, New York, USA) überführt.

Tabelle 7: Festgelegte Variablen zur Auswertung der EMERGE-Fälle

Variable	Angabe
Zeit bis Raumzuweisung	in Spielsekunden
Zeit bis erster Arzt-Patienten-Kontakt	in Spielsekunden
Prozedurales Vorgehen	als Buchstaben-/Zahlenkode
Vorgenommene Tätigkeiten	als Buchstaben-/Zahlenkode
Anzahl Anamnesefragen	absolut

Anzahl Laborparameter	absolut
Relevante Anamnesefragen	als Prozentscore (Anteil erreichter Punkte)
Relevante körperliche Untersuchung	dichotom
Relevante apparative Untersuchungen	als Prozentscore (Anteil erreichter Punkte)
Relevante Laborparameter	als Prozentscore (Anteil erreichter Punkte)
Relevante Therapiemaßnahmen	als Prozentscore (Anteil erreichter Punkte)
Erste und ggf. zweite gestellte Diagnose	als Buchstaben-/Zahlenkode
Richtige Diagnose	dichotom
Zeit bis erste Diagnose	in Spielsekunden
Zeit bis richtige Diagnose	in Spielsekunden
Gewählter Fallabschluss	als Buchstaben-/Zahlenkode
Richtiger Fallabschluss	dichotom
Zeit bis Fallabschluss	in Spielsekunden

2.5 Fokusgruppengespräche

2.5.1 Vorbereitung und Durchführung

Eine Woche nach Abschluss des Moduls 6.x wurden die Studierenden zur Teilnahme an Fokusgruppengesprächen zum Einsatz von EMERGE in der curricularen Lehre eingeladen. Hierzu wurde vom Studienleiter eine Einladung an die studentischen Email-Adressen aller Teilnehmer des Moduls 6.x versendet. Im Anhang dieser E-Mail befanden sich eine Teilnehmerinformation und eine Einwilligungserklärung zur Transkription und nachfolgenden Analyse der anonymisierten Gesprächsinhalte. Mittels der Terminplaner-Software „Doodle“ (Doodle AG, Zürich) wurde unter allen Interessierten eine elektronische Terminumfrage erstellt. Die endgültige Verteilung auf drei Termine zu je ca. 45 Minuten wurde den Studierenden anschließend per E-Mail mitgeteilt. Alle drei Termine fanden in den Räumlichkeiten des STPÄS statt.

Die Fokusgruppengespräche wurden in Form eines halb-standardisierten Interviews durchgeführt. Hierzu wurde im Vorfeld ein Interview-Leitfaden mit acht Leitfragen formuliert, die sich unter anderem mit der Organisation und der Gestaltung der EMERGE-Termine sowie dem subjektiven Lernerfolg beschäftigten. Zu jeder der acht Leitfragen wurden darüber hinaus bis zu fünf Unterfragen entwickelt (Fragenkatalog siehe Anhang 3). Der Interview-Leitfaden diente der Moderatorin der Fokusgruppengespräche, einer wissenschaftlichen Mitarbeiterin des Bereichs Medizindidaktik und Ausbildungsforschung der UMG zur Strukturierung der Diskussion. Die Fragenabfolge konnte dabei variiert und an den Gesprächsverlauf angepasst werden. Die Moderatorin informierte die Studierenden zu Beginn des Gesprächs erneut über Sinn und Ziel der Diskussion und verteilte die Teilnehmerinformation und die Einverständniserklärung. Außerdem klärte sie die Studierenden darüber auf, dass die Teilnahme freiwillig ist und die Diskussion elektronisch-

akustisch aufgezeichnet, schriftlich protokolliert und anschließend transkribiert wird, jedoch keinerlei persönliche Daten außer dem Geschlecht registriert werden.

Nach Unterzeichnung der Einverständniserklärung durch die Teilnehmer leitete die Moderatorin die Diskussion ein, indem sie die Studierenden fragte, was ihnen von den EMERGE-Terminen besonders in Erinnerung geblieben sei. Diese Frage war bewusst offen gestellt worden, um den Teilnehmern die Möglichkeit zu geben, selbstständig Diskussionsinhalte vorzuschlagen. Die Moderatorin regte die Studierenden im weiteren Gesprächsverlauf unter anderem dazu an, über die optimale Gestaltung eines EMERGE-Termins, erlebte Emotionen während der Nutzung der virtuellen Notaufnahme oder die Vor- und Nachteile von EMERGE im Vergleich zu anderen Lehrformen zu diskutieren.

Als Aufwandsentschädigung erhielt jeder Teilnehmer im Anschluss an die Fokusgruppendifkussion einen Büchergutschein im Wert von 20 Euro.

2.5.2 Transkription und Auswertung

Zur Transkription der elektronisch-akustischen Aufzeichnung wurden die Audiodateien abgespielt und wortwörtlich niedergeschrieben. Bei der Verschriftlichung wurden folgende Transkriptionsregeln angewandt: Es wurde vollständig wörtlich transkribiert, d. h. auch wiederholte Wörter, abgebrochene Sätze oder syntaktisch inkorrekte Satzformen wurden unverändert niedergeschrieben. Wortverschleifungen wurden geglättet und in korrektem Hochdeutsch wiedergegeben (z. B. „Das is´ schwierig.“ wurde zu „Das ist schwierig.“). Längere Pausen wurden durch drei Auslasspunkte in Klammern („...“) markiert; Einwürfe anderer Gesprächsteilnehmer sowie zustimmende Lautäußerungen, Mimik oder Gestik wurden in Klammern wiedergeben. Im Anschluss an die Verschriftlichung wurde in einem zweiten Durchgang das Transkript durch erneuten Vergleich mit der Audiodatei geprüft und korrigiert. Anschließend wurde die Audio-Datei gelöscht.

2.6 Statistische Analyse

Zunächst erfolgte eine deskriptive Analyse der vorliegenden Daten zur Charakterisierung der Stichprobe. Die statistische Analyse aller Daten erfolgte mit dem Statistikprogramm IBM SPSS Statistics für Windows, Version 24 (IBM Corp., Armonk, New York, USA). Das Signifikanzniveau wurde auf $0,05 = 5\%$ festgesetzt.

Zur Beantwortung der ersten Studienfrage wurden die Leistungen der Studierenden beider Lehrformen (POLEMA und EMERGE) in der formativen ‚Key Feature‘-Abschlussprüfung miteinander verglichen. Die statistische Auswertung erfolgte zunächst deskriptiv.

Anschließend wurde zum Vergleich der Leistungen beider Gruppen der Mann-Whitney-U-Test für die durchweg ordinalskalierten Variablen verwendet.

Zur Beantwortung der zweiten Studienfrage wurden die inhaltlichen und prozeduralen Ergebnisse der EMERGE-Abschlussprüfung a) zwischen der POLEMA und der EMERGE-Gruppe (Formatspezifität) und b) zwischen den EMERGE-Gruppen A, B und AB mit Bezug zum Ausmaß der individuellen Exposition während der Trainingsphase betrachtet (Fallspezifität). Im Anschluss an die deskriptive Analyse wurde zum Vergleich der Leistungen der Gruppen POLEMA und EMERGE erneut der Mann-Whitney-U-Test angewendet. Um die drei EMERGE-Gruppen miteinander vergleichen zu können, wurde der Kruskal-Wallis-H-Test durchgeführt; als Post-hoc-Test erfolgte anschließend in Teilbereichen mit signifikanten Gruppenunterschieden ein paarweiser Vergleich mittels Mann-Whitney-U-Test. Auf dieselbe Weise wurden zur Untersuchung der Fallspezifität auch die Ergebnisse der ‚*Key Feature*‘-Abschlussprüfung zwischen den drei EMERGE-Gruppen verglichen. Zudem erfolgte eine rein deskriptive, longitudinale Analyse aller vorliegenden Log-Dateien der 10 EMERGE-Seminartermine. Hierzu wurden die Ergebnisse der drei EMERGE-Gruppen tabellarisch gegenübergestellt.

Zur Beantwortung der dritten Studienfrage erfolgte eine deskriptive Auswertung und tabellarische Darstellung des strukturellen Vorgehens der EMERGE- und der POLEMA-Teilnehmer im Behandlungsablauf der vier EMERGE-Prüfungsfälle.

Zur Beantwortung der vierten und fünften Studienfrage wurden zunächst die Fragebögen deskriptiv ausgewertet und die Daten tabellarisch dargestellt. Aussagen, die nicht dichotom, sondern anhand einer (meist sechsstufigen) Likert-Skala zu bewerten waren, wurden als Mittelwert \pm Standardabweichung angegeben. Zudem wurde das Transkript der Fokusgruppengespräche von zwei voneinander unabhängigen Personen nach Mayring (2015) analysiert. Mithilfe der Software MAXQDA (VERBI Software GmbH, Berlin) wurden induktive Kategorien (sogenannte „Codes“) gebildet, die anschließend wesentlichen Textstellen des Transkripts zugeordnet wurden (sogenanntes „Coding“). Nach dem Vergleich der Ergebnisse beider Prüfer und der Zusammenfassung des Kategoriensystems durch die Bildung von Haupt- und Unterkategorien wurden die Themenschwerpunkte der Fokusgruppengespräche identifiziert. Zur Visualisierung wurde eine MindMap mithilfe der Software XMind 8 (XMind Ltd, Hongkong) erstellt. Die Ergebnisse wurden narrativ zusammengefasst.

3 Ergebnisse

3.1 Charakterisierung der Studiengruppe

Im Sommersemester 2016 nahmen insgesamt 153 Studierende am Modul 6.x teil (siehe Abbildung 2). Hiervon willigten 117 Studierende in die Teilnahme an der Studie ein. Es nahmen 35 Studierende an POLEMA teil, dessen Kapazität aufgrund der begrenzten Verfügbarkeit studentischer Tutoren beschränkt war. Da es in einem Fall zur Kontamination kam (durch einen Wechsel des Lehrformats am Ende der ersten Semesterwoche), bestand die Gruppe POLEMA letztendlich aus 34 Studienteilnehmern (52,9% weiblich). In der Gruppe EMERGE befanden sich 82 Studierende, davon 30 Studierende in der EMERGE-Gruppe A und jeweils 26 Studierende in den EMERGE-Gruppen B und AB. In Gruppe A kam es zu einer Kontamination und einem fehlenden Datensatz in der formativen Abschlussprüfung, in EMERGE-Gruppe B zu zwei fehlenden Datensätzen in der formativen Abschlussprüfung. Somit bestand die EMERGE-Gruppe letztendlich aus insgesamt 78 Studienteilnehmern (57,7% weiblich): 28 Teilnehmer (57,1% weiblich) in EMERGE-Gruppe A, 24 Teilnehmer (50% weiblich) in EMERGE-Gruppe B und 26 Teilnehmer (65,4% weiblich) in EMERGE-Gruppe AB. Es fanden sich weder signifikante Unterschiede hinsichtlich der Geschlechterverteilung in den Gruppen POLEMA vs. EMERGE ($p=0,643$) noch innerhalb der drei EMERGE-Gruppen ($p=0,549$).

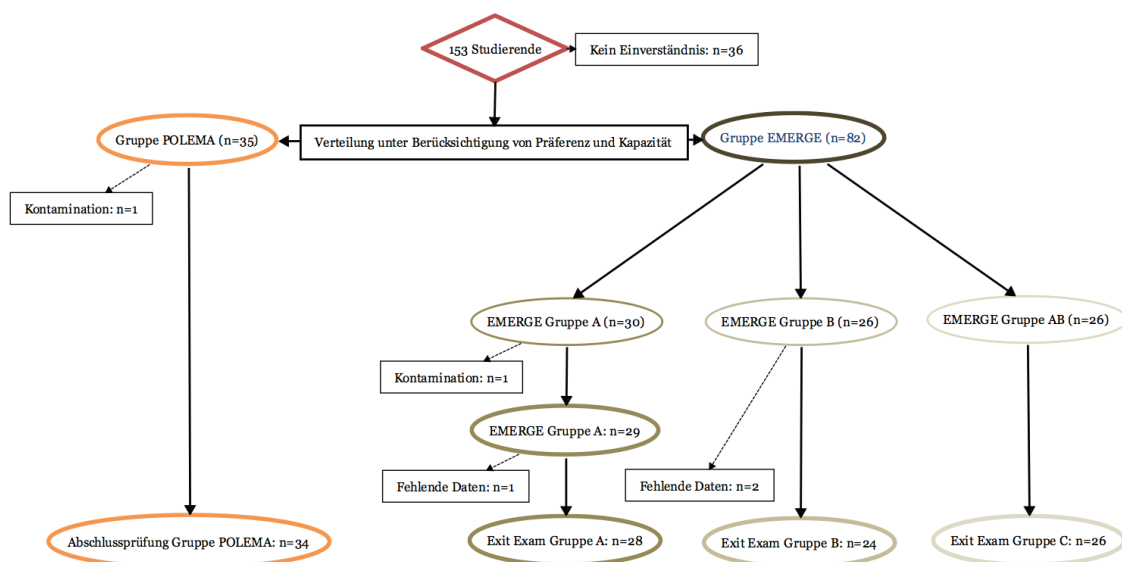


Abbildung 2: Zusammensetzung der Studiengruppe

Auch das durchschnittliche Alter unterschied sich nicht signifikant zwischen den Gruppen. Es betrug zu Beginn des Repetitionsmoduls 6.x in der Gruppe POLEMA $27,0 \pm 3,4$ Jahre

und in der Gruppe EMERGE $26,1 \pm 2,8$ Jahre ($p=0,323$). In der Gruppe EMERGE A betrug das durchschnittliche Alter $25,8 \pm 2,8$ Jahre, in der Gruppe EMERGE B $26,1 \pm 2,9$ Jahre und in der Gruppe EMERGE AB $26,4 \pm 2,8$ Jahre ($p=0,707$).

Bezüglich der Klausurleistungen im vorausgegangenen 9. Fachsemester unterschieden sich die Gruppen mit einer durchschnittlich erreichten Punktzahl von $87,7 \pm 5,1\%$ in der Gruppe POLEMA und $87,0 \pm 5,7\%$ in der Gruppe EMERGE nicht signifikant voneinander ($p=0,524$). Auch innerhalb der drei EMERGE-Gruppen (Gruppe A= $85,9 \pm 7,6\%$; Gruppe B= $87,7 \pm 4,6\%$; Gruppe AB= $87,6 \pm 4,3\%$) lag kein signifikanter Unterschied im Hinblick auf die Klausurleistungen des Vorsemesters vor ($p=0,947$).

3.2 Anwesenheit bei den Seminaren

Alle Studienteilnehmer haben an mindestens 80% der 10 Seminartermine im Rahmen von POLEMA bzw. EMERGE teilgenommen (exklusive der Abschlussprüfung in der Gruppe POLEMA). Die durchschnittliche Anwesenheit war in der Gruppe POLEMA mit $92,6 \pm 6,7\%$ signifikant höher als in der Gruppe EMERGE mit $87,1 \pm 8,4\%$; $p=0,001$ (EMERGE-Gruppe A = $87,1 \pm 8,1\%$; EMERGE-Gruppe B = $86,7 \pm 7,0\%$; EMERGE-Gruppe AB = $87,3 \pm 10,0\%$; $p=0,760$).

3.3 Studienfrage 1: Differentialdiagnostische/-therapeutische Kompetenzen

Zur Beantwortung der ersten Studienfrage („Erzielen Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, bessere bzw. schlechtere Ergebnisse im Hinblick auf differentialdiagnostische und –therapeutische Kompetenzen als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts lernen?“) wurden die Ergebnisse der ‚Key Feature‘-Abschlussprüfung herangezogen (siehe Abbildung 3). In dieser erzielten die EMERGE-Teilnehmer im Vergleich zu den POLEMA-Teilnehmern einen signifikant höheren Gesamtpunktwert ($61,8 \pm 14,5$ vs. $54,4 \pm 14,4$; $p=0,015$). Zudem zeigte sich ein signifikanter Vorteil der EMERGE-Teilnehmer auch in den beiden Einzelfällen „Herzinsuffizienz“ ($66,7 \pm 23,7$ vs. $47,8 \pm 26,4$; $p<0,001$) und „Hyponatriämie“ ($58,3 \pm 20,0$ vs. $44,9 \pm 22,0$; $p=0,003$). In drei weiteren Prüfungsfällen erzielten die EMERGE-Teilnehmer einen jeweils nicht signifikant höheren Punktwert als die POLEMA-Teilnehmer. Im Fall „Sarkoidose“ erzielte hingegen die Gruppe POLEMA einen nicht signifikant höheren Punktwert als die Gruppe EMERGE ($60,0 \pm 24,7$ vs. $57,7 \pm 24,3$; $p=0,536$).

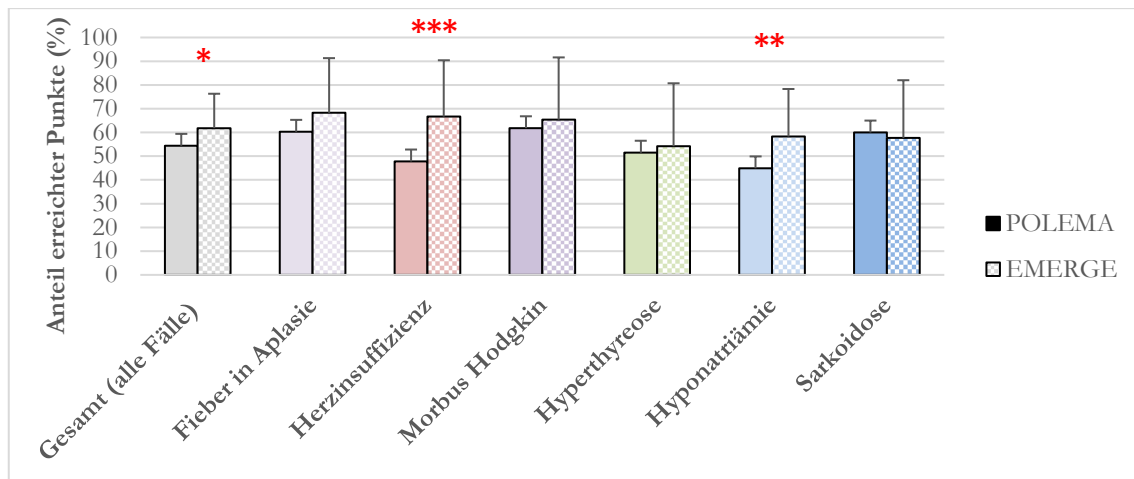


Abbildung 3: Ergebnisse der ‚Key Feature‘-Abschlussprüfung (Vergleich EMERGE und POLEMA).

Vergleich der Studiengruppen: * p<0,05; ** p<0,01; *** p<0,001 (Mann-Whitney-U-Test)

Es sei erwähnt, dass sich bezüglich Inhalten, welche während der Seminartermine zuvor sowohl in EMERGE als auch in POLEMA thematisiert worden waren (d. h. der Fälle „Sarkoidose“ und „Morbus Hodgkin“), kein signifikanter Unterschied im Hinblick auf die erreichte Gesamtpunktzahl zeigte (p=0,831) (siehe Abbildung 4).

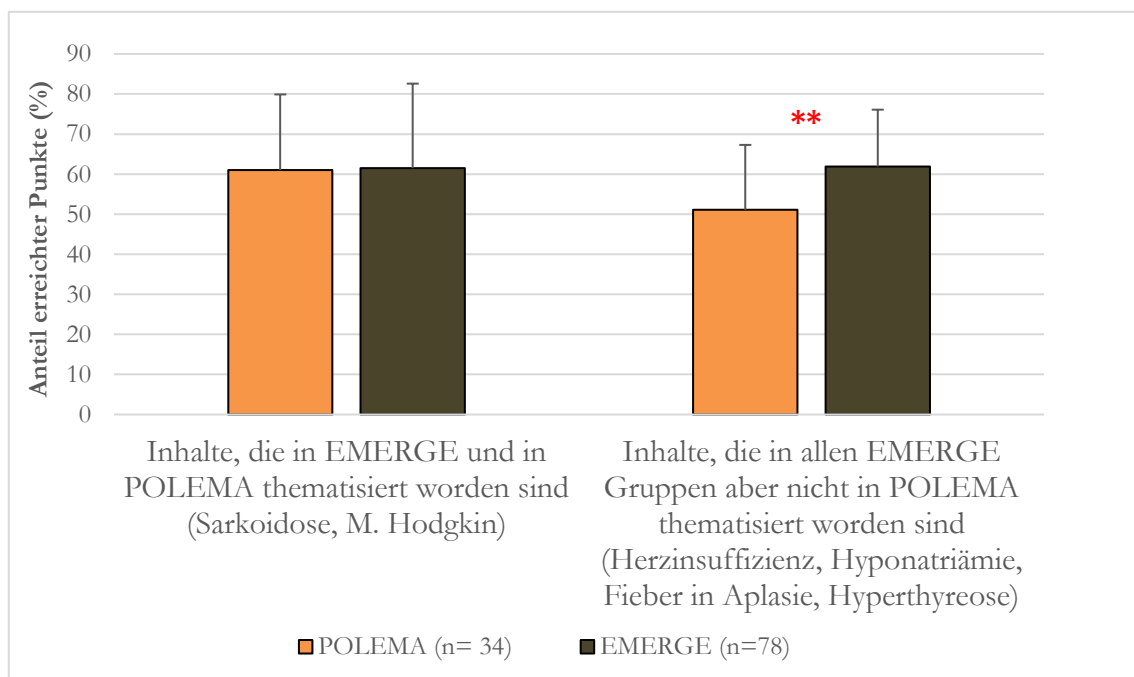


Abbildung 4: Ergebnisse der ‚Key Feature‘-Abschlussprüfung hinsichtlich der in den Lehrformen zuvor thematisierten Inhalte (Vergleich EMERGE und POLEMA).

Vergleich der Studiengruppen: * p<0,05; ** p<0,01; *** p<0,001 (Mann-Whitney-U-Test)

3.4 Studienfrage 2: Format- und/oder Fallspezifität der erlernten Kompetenzen

Die zweite Studienfrage („Inwieweit sind das im Rahmen eines problemorientierten Kleingruppenunterrichts oder EMERGE erlernte diagnostische und therapeutische Wissen sowie die Kompetenzen hinsichtlich Zeitmanagement und Priorisierung fallspezifisch und/ oder formatspezifisch?“) wurde auf Grundlage der Ergebnisse der EMERGE-Abschlussprüfung beantwortet.

3.4.1 Formatspezifität

In allen vier Fällen, die im Rahmen der EMERGE-Abschlussprüfung bearbeitet werden mussten, erzielten die EMERGE-Teilnehmer jeweils einen signifikant höheren Gesamtpunktwert als die POLEMA-Teilnehmer (siehe Tabelle 8). Im Hinblick auf inhaltliche Fähigkeiten (korrekte Diagnosestellung, adäquate Therapieeinleitung) ergab sich in sechs von acht Items ein signifikant höherer Punktwert.

Tabelle 8: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich der Globalbewertung und der inhaltlichen Fähigkeiten (Vergleich POLEMA und EMERGE)

ERGEBNISSE DER EMERGE-ABSCHLUSSPRÜFUNG I			
	POLEMA (n=34)	EMERGE (n=78)	p-Wert (Mann-Whitney-U-Test)
Globalbewertung (Anteil erreichter Punkte)			
NSTEMI	50,9 ± 13,2 (n=27)	66,4 ± 10,9 (n=74)	<0,001
Pankreatitis	47,6 ± 14,5 (n=26)	67,0 ± 10,9 (n=73)	<0,001
GI-Blutung	44,0 ± 9,1 (n=18)	60,5 ± 12,1 (n=70)	<0,001
Asthmaanfall	51,8 ± 9,9 (n=19)	60,2 ± 10,9 (n=73)	0,004
Korrekte Diagnose (dichotom)			
NSTEMI	70,4 (n=19)	97,3 (n=72)	<0,001
Pankreatitis	46,2 (n=12)	72,6 (n=53)	0,015
GI-Blutung	33,3 (n=6)	58,6 (n=41)	0,057
Asthmaanfall	84,2 (n=16)	97,3 (n=71)	0,026
Korrekte Therapie (Anteil erreichter Punkte)			
NSTEMI	24,7 ± 25,5 (n=27)	50,9 ± 35,9 (n=74)	0,001
Pankreatitis	15,4 ± 19,4 (n=26)	45,7 ± 29,1 (n=73)	<0,001
GI-Blutung	11,1 ± 32,3 (n=18)	14,3 ± 35,2 (n=70)	0,713
Asthmaanfall	42,1 ± 21,8 (n=19)	71,2 ± 24,4 (n=73)	<0,001

Weiterhin erzielten die EMERGE-Teilnehmer in acht von 19 Items zu prozeduralen Kompetenzen (korrekte Anamnese, korrekte körperliche Untersuchung, korrekte apparative Untersuchung, korrekte Laboranforderungen, korrekter Fallabschluss) einen signifikant höheren Punktwert als die POLEMA-Teilnehmer (siehe Tabelle 9). Während die

EMERGE-Teilnehmer in neun weiteren Items nicht signifikant höhere Punktwerte erreichten als die POLEMA-Teilnehmer, erzielten diese wiederum in zwei Items zum Fall „Asthmaanfall“ (korrekte körperliche Untersuchung, korrekter Fallabschluss) nicht signifikant höhere Punktwerte als die EMERGE-Teilnehmer.

Tabelle 9: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich prozeduraler Fähigkeiten (Vergleich POLEMA und EMERGE)

ERGEBNISSE DER EMERGE-ABSCHLUSSPRÜFUNG II			
	POLEMA (n=34)	EMERGE (n=78)	p-Wert (Mann-Whitney-U-Test)
Korrekte Anamnese (Anteil erreichter Punkte)			
NSTEMI	44,4 ± 17,8 (n=27)	54,5 ± 19,6 (n=74)	0,32
Pankreatitis	48,1 ± 23,7 (n=26)	57,5 ± 17,8 (n=73)	0,041
GI-Blutung	42,6 ± 20,8 (n=18)	52,6 ± 15,2 (n=70)	0,054
Asthmaanfall	50,0 ± 13,6 (n=19)	50,7 ± 16,5 (n=73)	0,988
Korrekte körperliche Untersuchung (dichotom)			
NSTEMI	81,5 (n=22)	97,3 (n=72)	0,006
Pankreatitis	96,2 (n=25)	98,6 (n=72)	0,443
GI-Blutung	66,7 (n=12)	94,3 (n=66)	0,001
Asthmaanfall	94,7 (n=18)	91,8 (n=67)	0,667
Korrekte apparative Untersuchung (Anteil erreichter Punkte)			
NSTEMI	38,3 ± 15,2 (n=27)	62,6 ± 20,6 (n=74)	<0,001
Pankreatitis	57,7 ± 50,4 (n=26)	94,5 ± 22,9 (n=73)	<0,001
GI-Blutung	5,6 ± 23,6 (n=18)	60,0 ± 49,3 (n=70)	<0,001
Asthmaanfall	§	§	§
Korrekte Laboranforderung (Anteil erreichter Punkte)			
NSTEMI	81,5 ± 36,2 (n=27)	91,9 ± 23,9 (n=74)	0,125
Pankreatitis	57,7 ± 30,6 (n=26)	88,6 ± 16,9 (n=72)	<0,001
GI-Blutung	66,7 ± 38,3 (n=18)	76,4 ± 26,5 (n=70)	0,407
Asthmaanfall	52,6 ± 51,3 (n=19)	69,9 ± 46,2 (n=73)	0,159
Korrekter Fallabschluss (dichotom)			
NSTEMI	77,8 (n=21)	82,4 (n=61)	0,598
Pankreatitis	53,8 (n=14)	58,9 (n=43)	0,656
GI-Blutung	66,7 (n=12)	90,0 (n=63)	0,013
Asthmaanfall	15,8 (n=3)	5,5 (n=4)	0,133

§ = nicht gewertet, da in diesem Fall keine apparativen Untersuchungen erwartet wurden

3.4.2 Fallspezifität

Um eine Fallspezifität der während der EMERGE-Seminare erlernten Kompetenzen zu untersuchen, wurden die Leistungen der drei EMERGE-Gruppen A, B und AB im Hinblick auf die vier (allen Studierenden unbekannt) EMERGE-Prüfungsfälle untersucht. Hierbei zeigte sich unter zwölf Items zu inhaltlichen Fähigkeiten (siehe Tabelle 10) und 19 Items zu prozeduralen Fähigkeiten (siehe Tabelle 11) jeweils lediglich ein signifikanter Leistungsunterschied zwischen den EMERGE-Gruppen.

Tabelle 10: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich der inhaltlichen Fähigkeiten (Vergleich EMERGE-Gruppen A, B und AB)

ERGEBNISSE DER EMERGE-ABSCHLUSSPRÜFUNG III				
	EMERGE-Gruppe A (n=28)	EMERGE-Gruppe B (n=24)	EMERGE-Gruppe AB (n=26)	p-Wert (Kruskal-Wallis-H-Test)
Globalbewertung (Anteil erreichter Punkte)				
NSTEMI	68,9 ± 12,2 (n=28)	66,0 ± 12,0 (n=21)	64,0 ± 7,6 (n=25)	0,231
Pankreatitis	66,5 ± 11,5 (n=25)	67,4 ± 9,2 (n=24)	67,2 ± 12,1 (n=24)	0,944
GI-Blutung	61,2 ± 12,2 (n=22)	61,5 ± 12,2 (n=24)	58,9 ± 12,2 (n=24)	0,683
Asthmaanfall	60,6 ± 11,0 (n=25)	57,4 ± 11,8 (n=24)	62,5 ± 9,7 (n=24)	0,296
Korrekte Diagnose (dichotom)				
NSTEMI	92,9 (n=26)	100,0 (n=21)	100,0 (n=25)	0,189
Pankreatitis	68,0 (n=17)	83,3 (n=20)	66,7 (n=16)	0,358
GI-Blutung	63,6 (n=14)	50,0 (n=12)	62,5 (n=15)	0,578
Asthmaanfall	100,0 (n=25)	91,7 (n=22)	100,0 (n=24)	0,126
Korrekte Therapie (Anteil erreichter Punkte)				
NSTEMI	56,0 ± 36,4 (n=28)	47,6 ± 20,9 (n=21)	48,0 ± 39,8 (n=25)	0,63
Pankreatitis	46,7 ± 30,4 (n=25)	41,7 ± 26,5 (n=24)	48,6 ± 31,1 (n=24)	0,087
GI-Blutung	45,5 ± 21,3 (n=22)	12,5 ± 33,8 (n=24)	25,0 ± 44,2 (n=24)	0,118
Asthmaanfall	76,0 ± 24,6 (n=25) ¹	59,7 ± 26,0 (n=24)	77,8 ± 18,8 (n=24) ²	0,026

¹ EMERGE-Gruppe A erreichte signifikant mehr Punkte als EMERGE-Gruppe B (asymptotische Signifikanz=0,032)

² EMERGE-Gruppe AB erreichte signifikant mehr Punkte als EMERGE-Gruppe B (asymptotische Signifikanz=0,012)

Tabelle 11: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich prozeduraler Fähigkeiten (Vergleich EMERGE-Gruppen A, B und AB)

ERGEBNISSE DER EMERGE-ABSCHLUSSPRÜFUNG IV				
	EMERGE-Gruppe A (n=28)	EMERGE-Gruppe B (n=24)	EMERGE-Gruppe AB (n=26)	p-Wert (Kruskal-Wallis-H-Test)
Korrekte Anamnese (Anteil erreichter Punkte)				
NSTEMI	56,7 ± 20,5 (n=28)	57,7 ± 24,5 (n=21)	49,0 ± 11,9 (n=25)	0,414
Pankreatitis	55,3 ± 17,8 (n=25)	59,0 ± 19,0 (n=24)	58,3 ± 17,0 (n=24)	0,691

GI-Blutung	51,5 ± 14,5 (n=22)	54,9 ± 14,3 (n=24)	51,4 ± 17,0 (n=24)	0,477
Asthmaanfall	47,3 ± 17,1 (n=25)	53,5 ± 17,0 (n=24)	51,4 ± 15,5 (n=24)	0,293
Korrekte körperliche Untersuchung (dichotom)				
NSTEMI	96,4 (n=27)	95,2 (n=20)	100,0 (n=25)	0,577
Pankreatitis	100,0 (n=25)	100,0 (n=24)	95,8 (n=23)	0,36
GI-Blutung	90,9 (n=20)	95,8 (n=23)	95,8 (n=23)	0,716
Asthmaanfall	96,0 (n=24)	87,5 (n=21)	91,7 (n=22)	0,561
Korrekte apparative Untersuchung (Anteil erreichter Punkte)				
NSTEMI	65,5 ± 21,2 (n=28)	58,7 ± 18,0 (n=21)	62,7 ± 22,2 (n=25)	0,552
Pankreatitis	100,0 ± 0,0 (n=25)	91,7 ± 28,2 (n=24)	91,7 ± 23,2 (n=24)	0,337
GI-Blutung	59,1 ± 50,3 (n=22)	62,5 ± 49,5 (n=24)	58,3 ± 50,4 (n=24)	0,953
Asthmaanfall	§	§	§	§
Korrekte Laboranforderung (Anteil erreichter Punkte)				
NSTEMI	94,0 ± 22,3 (n=28)	85,7 ± 30,9 (n=21)	94,7 ± 18,5 (n=25)	0,362
Pankreatitis	92,0 ± 14,5 (n=25)	87,5 ± 16,5 (n=24)	86,1 ± 19,5 (n=24)	0,487
GI-Blutung	88,6 ± 21,4 (n=22) ¹	77,1 ± 25,4 (n=24)	64,6 ± 27,5 (n=24)	0,01
Asthmaanfall	68,0 ± 47,6 (n=25)	66,7 ± 48,2 (n=24)	75,0 ± 44,2 (n=24)	0,798
Korrektur Fallabschluss (dichotom)				
NSTEMI	89,3 (n=25)	85,7 (n=18)	72,0 (n=18)	0,234
Pankreatitis	48,0 (n=12)	62,5 (n=15)	66,7 (n=16)	0,382
GI-Blutung	90,9 (n=20)	95,8 (n=23)	83,3 (n=20)	0,353
Asthmaanfall	12,0 (n=3)	0,0 (n=0)	4,2 (n=1)	0,176

¹ EMERGE-Gruppe A erreichte signifikant mehr Punkte als EMERGE-Gruppe AB (asymptotische Signifikanz=0,003)
 § = nicht gewertet, da in diesem Fall keine apparativen Untersuchungen erwartet wurden

In der ‚Key Feature‘-Abschlussprüfung (siehe Tabelle 12) unterschieden sich die drei EMERGE-Gruppen A, B und AB nicht hinsichtlich der erreichten Gesamtpunktezahl, in welche die Leistungen aller sechs Fälle eingingen ($p=0,714$). Auch im Hinblick auf die sechs Einzelfälle gab es keine signifikanten Leistungsunterschiede. Nichtsdestoweniger hat die höchste Punktezahl für einen Fall jeweils die Gruppe erzielt, welche den Inhalten des jeweiligen Fachgebiets gegenüber während der EMERGE-Seminare (am stärksten) exponiert worden war. Entsprechend wurden die höchsten Punktwerte in den kardiologischen/pulmologischen und gastroenterologischen/endokrinologischen Fällen, d. h. Herzinsuffizienz, Hyponatriämie und Sarkoidose, von EMERGE-Gruppe A erzielt. Insgesamt erreichte die Gruppe für diese drei Fälle mit $62,8 \pm 16,4\%$ ein nicht signifikant höheres Ergebnis als die EMERGE-Gruppen B ($59,7 \pm 15,5\%$) und AB ($59,9 \pm 16,5\%$; $p=0,618$) (Abbildung 5). Im Gegensatz dazu erzielte EMERGE-Gruppe B die jeweils höchsten Punktwerte in den nephrologischen/rheumatologischen und hämatologischen/onkologischen Fällen, d. h. Fieber in Aplasie, Morbus Hodgkin und Hyperthyreose. Mit einer prozentualen Gesamtleistung von $68,4 \pm 12,3\%$ hinsichtlich dieser

drei Fälle ist auch dieses Ergebnis nicht signifikant besser als in den EMERGE-Gruppen A ($60,4 \pm 20,4\%$) und AB ($59,6 \pm 16,8$; $p=0,216$).

Tabelle 12: Ergebnisse der ‚Key Feature‘-Abschlussprüfung (Vergleich EMERGE-Gruppen A, B und AB)

ERGEBNISSE ‚KEY FEATURE‘-ABSCHLUSSPRÜFUNG				
	EMERGE-Gruppe A (n=28)	EMERGE-Gruppe B (n=24)	EMERGE-Gruppe AB (n=26)	p-Wert (Kruskal-Wallis-H-Test)
Globalbewertung (Anteil erreichter Punkte)				
Gesamt (alle Fälle)	61,6 ± 16,3	64,1 ± 12,3	59,8 ± 14,7	0,714
Fieber in Aplasie	62,5 ± 26,8	71,9 ± 18,5	71,2 ± 22,0	0,374
Herzinsuffizienz	67,9 ± 23,4	66,7 ± 19,0	65,4 ± 28,4	0,954
Morbus Hodgkin	66,1 ± 27,4	71,9 ± 19,9	58,7 ± 29,1	0,276
Hyperthyreose	52,7 ± 25,8	61,5 ± 28,5	49,0 ± 25,0	0,259
Hyponatriämie	62,5 ± 24,1	55,2 ± 20,8	56,7 ± 15,1	0,212
Sarkoidose	58,0 ± 22,6	57,3 ± 29,0	57,7 ± 22,1	0,98

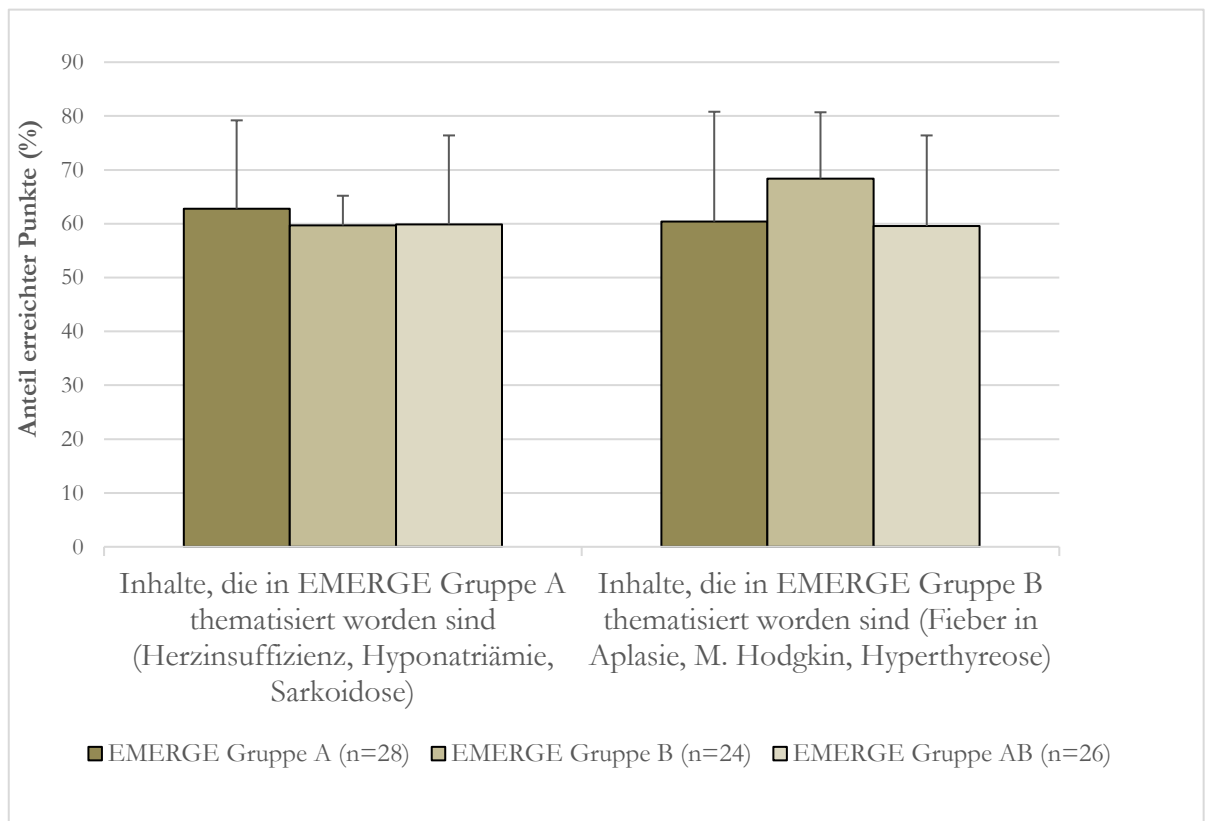


Abbildung 5: Ergebnisse der ‚Key Feature‘-Abschlussprüfung im Hinblick auf die inhaltlichen Schwerpunkte der drei EMERGE-Gruppen (Vergleich EMERGE A, B und AB)

3.5 Studienfrage 3: Prozedurales Vorgehen

Zur Beantwortung der dritten Studienfrage („Geben Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, im Behandlungsablauf von virtuellen Patienten strukturell anders vor als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts diskutieren?“) wurde das strukturelle Vorgehen der EMERGE- und der POLEMA-Teilnehmer im Behandlungsablauf der vier EMERGE-Prüfungsfälle verglichen (siehe Tabelle 13).

Im Hinblick auf das allgemeine Patientenmanagement war der Anteil abgeschlossener Fälle (d. h. Prüfungsfälle, die durch Verlegung auf eine andere Station oder Entlassung nach Hause im Spiel „abgeschlossen“ wurden) für alle vier Patientengeschichten in der EMERGE-Gruppe höher als in der POLEMA-Gruppe. Den Fall „Asthmaanfall“ schlossen beispielsweise 93,6% (n=73) der Studierenden der EMERGE-Gruppe und nur 55,9% (n=19) der Studierenden der POLEMA-Gruppe ab. Ähnliche Zahlen ergaben sich für den Fall „Untere gastrointestinale Blutung“, den 89,7% (n=70) der EMERGE-Teilnehmer und 52,9% (n=18) der POLEMA-Teilnehmer abschlossen.

Bezüglich des strukturellen Vorgehens lässt sich Folgendes zusammenfassen: In beiden Gruppen wurde in allen vier Fällen am häufigsten mit der Anamnese begonnen (Gruppe POLEMA: Minimum=66,7% vs. Maximum=92,3%; Gruppe EMERGE: Minimum=66,2% vs. Maximum=94,6%). Die POLEMA-Teilnehmer führten anschließend am häufigsten zunächst eine körperliche Untersuchung gefolgt von Labor/apparativen Untersuchungen durch. Im Gegensatz dazu entschieden sich die EMERGE-Teilnehmer im Anschluss an die Anamnese am häufigsten dazu, zunächst Labor/apparative Untersuchungen anzufordern und erst an dritter Stelle eine körperliche Untersuchung durchzuführen. Am deutlichsten war dieser Unterschied bei der Bearbeitung des Falles „Pankreatitis“ zu sehen: Während in der Gruppe POLEMA 80,8% das erstgenannte Vorgehen wählten, taten dies nur 28,8% der EMERGE-Teilnehmer. Demgegenüber entschieden sich in der EMERGE-Gruppe 65,8% für das zweitgenannte Vorgehen (Anamnese; Labor/apparative Untersuchungen; körperliche Untersuchung).

Im Hinblick auf den Fall „NSTEMI“ fiel in beiden Gruppen eine Priorisierung von Labor/apparativen Untersuchungen auf. So haben 31,1% der EMERGE-Teilnehmer (und 3,7% der POLEMA-Teilnehmer) als erstes Labor/apparative Untersuchungen angefordert. In den drei übrigen Fällen gingen jeweils weniger als 3% der EMERGE-Teilnehmer (und kein einziger POLEMA-Teilnehmer) auf diese Weise vor. Auch 33,3% der POLEMA-Teilnehmer entschieden sich im Fall „NSTEMI“ dazu, Labor/apparative Untersuchungen

zeitlich früher durchzuführen und gingen entsprechend nach der zweitgenannten Weise vor (Anamnese; Labor/apparative Untersuchungen; körperliche Untersuchung). Dieses Vorgehen wurde in den drei übrigen Fällen nur von jeweils weniger als 17% der POLEMA-Teilnehmer gewählt.

Tabelle 13: Ergebnisse der EMERGE-Abschlussprüfung hinsichtlich des prozeduralen Vorgehens (Vergleich POLEMA und EMERGE)

PROZEDURALES VORGEHEN IN DER EMERGE-ABSCHLUSSPRÜFUNG								
	NSTEMI		PANKREATITIS		UNTERE GI-BLUTUNG		ASTHMAANFALL	
	POLEMA (n=34)	EMERGE (n=78)	POLEMA (n=34)	EMERGE (n=78)	POLEMA (n=34)	EMERGE (n=78)	POLEMA (n=34)	EMERGE (n=78)
Management								
Der Spieler hat den Fall abgeschlossen.	79,4 (n=27)	94,9 (n=74)	76,5 (n=26)	93,6 (n=73)	52,9 (n=18)	89,7 (n=70)	55,9 (n=19)	93,6 (n=73)
Der Spieler hat den Fall nicht abgeschlossen aber mind. eine Aktivität vorgenommen.	20,6 (n=7)	5,1 (n=4)	23,5 (n=8)	6,4 (n=5)	47,1 (n=16)	10,3 (n=8)	44,1 (n=15)	6,4 (n=5)
Prozedurales Vorgehen (nur unter Studierenden, die den jeweiligen Fall abgeschlossen haben)								
1. Anamnese; 2. Körperliche Untersuchung; 3. Labor/apparative Untersuchungen	44,4 (n=12)	17,6 (n=13)	80,8 (n=21)	28,8 (n=21)	50,0 (n=9)	18,6 (n=13)	52,6 (n=10)	31,5 (n=23)
1. Anamnese; 2. Labor/apparative Untersuchungen; 3. Körperliche Untersuchung	33,3 (n=9)	48,6 (n=36)	11,5 (n=3)	65,8 (n=48)	16,7 (n=3)	67,1 (n=47)	15,8 (n=3)	39,7 (n=29)
1. Körperliche Untersuchung; 2. Anamnese; 3. Labor/apparative Untersuchungen	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)	0,0 (n=0)
1. Labor/apparative Untersuchungen; 2. Anamnese; 3. Körperliche Untersuchung	3,7 (n=1)	31,1 (n=23)	0,0 (n=0)	2,7 (n=2)	0,0 (n=0)	2,9 (n=2)	0,0 (n=0)	2,7 (n=2)
Sonstige	18,5 (n=5)	2,7 (n=2)	7,7 (n=2)	2,7 (n=2)	33,3 (n=6)	11,4 (n=8)	31,6 (n=6)	26,0 (n=19)

3.6 Studienfragen 4 und 5: Ergebnisse der Fragebögen und der Fokusgruppengespräche

Zur Beantwortung der vierten Studienfrage („Empfinden Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, bei der gleichzeitigen Behandlung von mehreren virtuellen Patienten weniger Stress als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts diskutieren?“) und der fünften Studienfrage („Halten Studierende die Computersimulation EMERGE für ein geeignetes Lernformat zum Erwerb ärztlicher Kompetenzen?“) wurden die Fragebögen und die Fokusgruppengespräche ausgewertet.

3.6.1 Ergebnisse der Fragebögen

Die Ergebnisse der papierbasierten Studierendenbefragung mit insgesamt 142 Teilnehmern (darunter 96 EMERGE-Teilnehmer und 46 POLEMA-Teilnehmer) sind in Tabelle 14 dargestellt. Die Gruppen unterschieden sich weder hinsichtlich des Geschlechterverhältnisses ($p=0,959$) noch hinsichtlich des durchschnittlichen Alters ($p=0,954$). Ebenso zeigten sich keine Unterschiede bezüglich der individuellen Vorerfahrung im Rettungswesen. Im Hinblick auf die Nutzung digitaler Medien gaben beide Gruppen gute Kenntnisse im Umgang mit Computern an. Allerdings hatten nur 10,9% der POLEMA-Teilnehmer und 15,5% der EMERGE-Teilnehmer zuvor jemals ein anderes ‚Serious Game‘ als EMERGE genutzt ($p=0,506$). Die Zustimmung zur Aussage „Mir fällt es meist leicht, zu erlernen, wie man ein neues Computerprogramm verwendet“ lag unter den EMERGE-Teilnehmern ($2,2 \pm 1,2$) signifikant höher als unter den POLEMA-Teilnehmern ($2,9 \pm 1,5$; $p=0,001$). Hinsichtlich des Stressempfindens bei der gleichzeitigen Behandlung mehrerer virtueller Patienten während der EMERGE-Abschlussprüfung gaben beide Gruppen ein mäßiges Stresslevel an. Nichtsdestoweniger fühlten sich die POLEMA-Teilnehmer signifikant gestresster als die EMERGE-Teilnehmer ($3,3 \pm 1,6$ vs. $4,3 \pm 1,4$; $p=0,001$).

Tabelle 14: Ergebnisse der Fragebögen hinsichtlich soziodemographischer Daten, Vorerfahrung im Rettungswesen, Nutzung digitaler Medien und Stressempfinden (Vergleich POLEMA und EMERGE)

ERGEBNISSE DER FRAGEBÖGEN I		
	POLEMA (n=46)	EMERGE (n=96)
Soziodemographische Daten		
Weiblich	60,9% (n=28)	59,8% (n=58)
Alter (in Jahren)	$26,5 \pm 2,86$	$26,67 \pm 3,48$

Vorerfahrung im Rettungswesen			
"Haben Sie schon einmal in einer Notaufnahme gearbeitet?" (Anteil "ja")		39,1% (n=18)	33,0% (n=32)
"Haben Sie schon einmal im Rettungsdienst gearbeitet?" (Anteil "ja")		13,0% (n=6)	18,6% (n=18)
Nutzung digitaler Medien			
"Wie schätzen Sie ihre Kenntnisse im Umgang mit Computern ein?" [§]	Gar keine	0% (n=0)	1,0% (n=1)
	Anfänger/in	28,3% (n=13)	30,2% (n=29)
	Fortgeschrittene/r	65,2% (n=30)	59,4% (n=57)
	Expertin/Experte	6,5% (n=3)	9,4% (n=9)
"Wie viele Stunden verbringen Sie pro Tag insgesamt am Computer?" [§]	<1h pro Tag	13,0% (n=6)	11,6% (n=11)
	1-2h pro Tag	21,7% (n=10)	34,7% (n=33)
	2-3h pro Tag	32,6% (n=15)	25,3% (n=24)
	3-4h pro Tag	19,6% (n=9)	10,5% (n=10)
	4-5h pro Tag	6,5% (n=3)	14,7% (n=14)
	>5h pro Tag	6,5% (n=3)	3,2% (n=3)
"Wie regelmäßig spielen Sie Computerspiele?" [§]	Nie	57,8% (n=26)	48,4% (n=46)
	1x pro Jahr	15,6% (n=7)	25,3% (n=24)
	1x pro Monat	15,6% (n=7)	13,7% (n=13)
	1x pro Woche	4,4% (n=2)	11,6% (n=11)
	Jeden Tag	6,5% (n=3)	1,1% (n=1)
"Haben Sie schon einmal ein anderes Lernspiel als EMERGE genutzt?" (Anteil "ja")		10,9% (n=5)	15,5% (n=15)
"Beim Arbeiten mit Computern habe ich Spaß."		3,1 ± 1,6	2,4 ± 1,1 *
"Ich finde, dass Computer beim Lernen hilfreich sind."		2,4 ± 1,2	1,9 ± 0,9 *
"Mir fällt es meist leicht, zu erlernen, wie man ein neues Computerprogramm verwendet."		2,9 ± 1,5	2,2 ± 1,2 **
Stressempfinden			
"Beim Spielen der heutigen EMERGE-Abschlussstzung habe ich mich gestresst gefühlt."		3,3 ± 1,6	4,3 ± 1,4 **
Legende: Eventuelle Inkonsistenzen in den Summen sind auf eine jeweils unterschiedliche Anzahl fehlender Werten pro Item zurückzuführen. Alle Aussagen waren entweder dichotom oder anhand einer sechsstufigen Likertskala (1 = „trifft voll zu“; 6 = „trifft überhaupt nicht zu“) zu beantworten. Ausnahmen stellen die mit „§“ gekennzeichneten Aussagen dar, bei denen die Häufigkeitsverteilung der Antwortmöglichkeiten angegeben ist. Vergleich der Studiengruppen: * p<0,05; ** p<0,01; *** p<0,001			

Die EMERGE-Teilnehmer beantworteten zudem 8 Fragen bezüglich ihrer Motivation zur Teilnahme an EMERGE und 6 Fragen bezüglich der inhaltlichen Bewertung der virtuellen Notaufnahme (siehe Tabelle 15).

Die meisten Studierenden entschieden sich zur Teilnahme an EMERGE, da sie sich für den Einsatz von Lernspielen im Medizinstudium interessieren ($2,8 \pm 1,8$) und lieber selbstständig Fälle bearbeiten, als diese in einer Kleingruppe (wie in POLEMA) zu besprechen ($2,8 \pm 1,8$). Zudem stimmten die Studierenden den Aussagen zu, EMERGE zum Training differentialdiagnostischer Fertigkeiten, zur Erweiterung von therapeutischem Wissen und zur besseren Einschätzung der Dringlichkeit klinischer Situationen gewählt zu haben. Des

Weiteren gaben die EMERGE-Teilnehmer selbsteinschätzend an, ihre differentialdiagnostischen und therapeutischen Fähigkeiten auch tatsächlich verbessert zu haben. So stimmten sie unter anderem den Aussagen zu, durch das Spielen von EMERGE die Dringlichkeit klinischer Situationen und die Wertigkeit verschiedener diagnostischer Schritte besser einschätzen zu können. Insgesamt schätzten sie EMERGE als sehr gut geeignet zur Vermittlung ärztlicher Kompetenzen im Medizinstudium ein ($1,74 \pm 0,8$).

Tabelle 15: Ergebnisse der Fragebögen hinsichtlich der Motivation zur Teilnahme an EMERGE und der inhaltlichen Bewertung der virtuellen Notaufnahme (nur EMERGE)

ERGEBNISSE DER FRAGEBÖGEN II	
	EMERGE (n=96)
Motivation: "Ich habe mich zur Teilnahme an EMERGE entschieden..."	
„..., weil ich mich für den Einsatz von Lernspielen im Medizinstudium interessiere.“	2,8 ± 1,8
„..., weil ich lieber selbstständig Fälle bearbeite, als diese in einer Kleingruppe zu besprechen.“	2,8 ± 1,8
„..., weil ich dadurch einen geringeren Arbeitsaufwand als bei POLEMA erwartet habe.“	4,7 ± 1,5
„..., weil ich der EMERGE-Gruppe zugelost wurde.“	4,4 ± 2,2
„..., weil ich gerne am Computer arbeite.“	3,4 ± 1,5
„..., um differentialdiagnostische Fertigkeiten zu trainieren.“	2,3 ± 1,5
„..., um mein therapeutisches Wissen zu erweitern.“	2,4 ± 1,6
„..., um zu lernen, die Dringlichkeit klinischer Situationen richtig einzuschätzen.“	2,3 ± 1,6
Inhaltliche Bewertung	
„Ich fühlte mich bei der gleichzeitigen Behandlung mehrerer Patienten während der EMERGE-Übungssitzungen gestresst.“	3,4 ± 1,3
„Ich halte EMERGE zur Vermittlung ärztlicher Kompetenzen im Medizinstudium geeignet.“	1,7 ± 0,8
„Durch das Spielen von EMERGE habe ich meine Kompetenz gesteigert, die zu seinem Symptomkomplex gehörenden Differentialdiagnosen zu kennen.“	2,1 ± 1,0
„Durch das Spielen von EMERGE habe ich gelernt, die Dringlichkeit einer klinischen Situation abzuschätzen.“	2,4 ± 1,2
„Durch das Spielen von EMERGE habe ich gelernt, die Wertigkeit verschiedener diagnostischer Schritte einschätzen zu können.“	2,5 ± 1,2
„Durch das Spielen von EMERGE habe ich gelernt, die bei einer Diagnose relevanten therapeutischen Schritte einzuleiten.“	2,3 ± 1,2
Legende: Eventuelle Inkonsistenzen in den Summen sind auf eine jeweils unterschiedliche Anzahl fehlender Werte pro Item zurückzuführen. Alle Aussagen waren entweder dichotom oder anhand einer sechsstufigen Likertskala (1= „trifft voll zu“, 6 = „trifft überhaupt nicht zu“) zu beantworten.	

3.6.2 Ergebnisse der Fokusgruppengespräche

3.6.2.1 Teilnehmer

An den drei Fokusgruppengesprächen im Sommersemester 2016 nahmen 13 weibliche und 5 männliche Studierende ($n = 18$) teil. Ein Gesprächsteilnehmer hatte während des Repetitionsmoduls 6.x die POLEMA-Seminare besucht; die 17 weiteren Teilnehmer stammten aus der EMERGE-Gruppe.

3.6.2.2 Inhaltliche Ergebnisse

Im Rahmen der qualitativen Inhaltsanalyse nach Mayring (2015) wurden insgesamt 179 induktive Codes gebildet und 588 Codings vorgenommen. Im Zuge der Zusammenfassung des Kategoriensystems wurden die folgenden Hauptkategorien gebildet, welche die fünf Themenschwerpunkte der Fokusgruppengespräche widerspiegeln: I. EMERGE im Vergleich zu anderen Lehrformen; II. Inhaltliche Gestaltung; III. Organisation des Lehrformats; IV. Emotionen; V. Technische Aspekte.

Innerhalb dieser Hauptkategorien wurde durch die Bildung von Unterkategorien eine weitere Strukturierung der thematischen Inhalte vorgenommen. Die zum Zwecke der Visualisierung und Kommunikation erstellte MindMap umfasst alle Codes und befindet sich im Anhang dieser Arbeit (siehe Anhang 4). Im Folgenden werden die o.g. fünf thematischen Schwerpunkte vorgestellt. Hierbei erfolgt eine Beschränkung auf Aspekte mit mindestens sieben Codings (siehe Tabelle 16 bis Tabelle 20). Die aufgeführten Zitate dienen der Repräsentation charakteristischer Meinungen der Gesprächsteilnehmer bezüglich eines Gesprächsinhalts.

I. EMERGE im Vergleich zu anderen Lehrformen

Das Thema „EMERGE im Vergleich zu anderen Lehrformen“ stellte mit rund 47% aller Codings (275 von 588 Codings) den inhaltlichen Schwerpunkt der Fokusgruppendifkussionen dar (siehe Tabelle 16). Die Gesprächsteilnehmer nannten diesbezüglich sowohl Vor- als auch Nachteile.

Als größter Vorteil wurde das Training der Organisationsfähigkeit von den Gesprächsteilnehmern hervorgehoben (86 Codings). Durch EMERGE haben sie nach eigenen Angaben ihre Fähigkeit verbessert, Priorisierungen vorzunehmen und die Dringlichkeit klinischer Situationen einzuschätzen. Darüber hinaus zählten sie eine verbesserte Selbstorganisation zu ihren größten Lernerfolgen. Eine Studentin beantwortete

die Frage, was sie für ihren späteren Berufsalltag habe lernen können, wie folgt: *„Ja, Arbeitsmanagement, Zeitmanagement, Priorisierung. Wenn der Notfall kommt, dann musst du halt die Frau mit Kopfschmerzen warten lassen.“* (Fokusgruppengespräch (FG) Nummer (Nr.) 3, Zeile 124). Eine weitere Studentin im zweiten Fokusgruppengespräch war der gleichen Meinung: *„(...) einfach, dass ich mir überlege: ‚Was ist jetzt das Wichtigste, wo brennt es, wo muss ich als Erstes hin?‘.“* (FG Nr. 2, Zeile 23). Die Studierenden berichteten außerdem über die Entwicklung von Strategien zur Optimierung ihrer Arbeitsabläufe. Insbesondere das Notieren von Stichpunkten bezüglich Patientendaten und Behandlungsabläufen auf einem Notizzettel wurde als hilfreich empfunden. Besonders betont wurde von den Gesprächsteilnehmern darüber hinaus, dass EMERGE selbstständiges Arbeiten (49 Codings) fordert und fördert. Es bestehe ein Handlungs- und Entscheidungsdruck, welcher von den Studierenden als sehr positiv wahrgenommen wurde. Auch das Erleben von direkten Konsequenzen des eigenen Handelns auf den Gesundheitszustand des Patienten wurde positiv hervorgehoben, da Behandlungsfehler auf diese Weise in Erinnerung blieben und aus ihnen gelernt wurde. Es sei ein großer Vorteil, *„dass man wirklich seine Entscheidung für sich treffen muss und dann gucken muss, war das gut oder war das eher nicht so gut.“* (FG Nr. 3, Zeile 38, männlich). EMERGE wurde von den Studierenden als realistisches und praxisnahes Szenario wahrgenommen, welches eine gute Vorbereitung auf den späteren Berufsalltag darstellt. So beantwortete eine Studentin die Frage, ob sie etwas für ihren späteren Berufsalltag habe lernen können, mit der Aussage: *„Auf jeden Fall! Man hat ja wirklich eigentlich jetzt ein bisschen auf der TN gearbeitet.“* (FG Nr. 3, Zeile 118). Die Gesprächsteilnehmer vermerkten weiterhin einen Lernzuwachs bezüglich Differentialdiagnosen und Therapieverfahren sowie der Aneignung von Behandlungsalgorithmen.

In Hinblick auf die Nachteile übten die Studierenden deutliche Kritik am Feedback (47 Codings), welches sie nach Abschluss eines Patientenfalls im Archivmenü des Spiels einsehen konnten. *„Was ich schade fand, war, dass das Feedback, was man durch das Spiel erhalten hat, einem so gar nicht geholfen hat, einzuschätzen, was man am besten nochmal nachschauen konnte. Also ich fand das Feedback unglaublich schlecht an dem Spiel und das war, glaube ich, auch der größte Kritikpunkt (...) ([eine Teilnehmerin] wirft ein: Schließe ich mich an!).“* (FG Nr. 2, Zeile 39), äußerte sich ein männlicher Gesprächsteilnehmer. Aufgrund der knappen Rückmeldung und ihres begrenzten Informationsgehalts sind Fehler teilweise unbemerkt geblieben: *„Im schlechtesten Fall wusstest du überhaupt nicht, dass du etwas falsch gemacht hast.“* (FG Nr. 1, Zeile 86, weiblich).

Tabelle 16: Ergebnisse der Fokusgruppengespräche – EMERGE im Vergleich zu anderen Lehrformen

Code	Codings	Interviews
I. EMERGE im Vergleich zu anderen Lehrformen	275	3
I.1 Vorteile	201	3
I.1.1 Training der Organisationsfähigkeit	86	3
I.1.1.1 Lernerfolge	52	3
I.1.1.1.1 Vornehmen von Priorisierungen	27	3
I.1.1.1.1.1 Priorisierungen bzgl. verschiedener Patienten	14	3
I.1.1.1.1.1.a Einschätzung der Dringlichkeit klinischer Situationen	12	3
I.1.1.2 Selbstorganisation	20	2
I.1.1.2 Organisationsformen	34	3
I.1.1.2.1 Abfolge der Patientenbehandlung	15	3
I.1.1.2.2 Optimierung der Arbeitsabläufe	14	3
I.1.1.2.2.1 Mitschrift auf Notizzettel	11	3
I.1.2 Selbstständiges Arbeiten	49	3
I.1.2.1 Erleben von Konsequenzen des eigenen Handelns	17	2
I.1.2.1.1 Fehler sind im Gedächtnis geblieben	14	3
I.1.2.1.1.1 Lernen aus Fehlern	12	3
I.1.2.1.2 Handlungs- und Entscheidungsdruck	8	3
I.1.3 Realistisches Szenario	25	3
I.1.3.1 Gute Vorbereitung auf späteren Berufsalltag	10	2
I.1.4 Lernen von Algorithmen und Routinen	13	3
I.1.5 Lernzuwachs bzgl. Differentialdiagnosen und Therapie	7	2
I.2 Nachteile	74	3
I.2.1 Feedback	47	3
I.2.1.1 Schlechtes Feedback im Spiel	36	3
I.2.1.1.1 Fehler sind teilweise unbemerkt geblieben	14	3
I.2.1.2 Wunsch nach besserem Feedback	8	2
I.2.2 Technische Grenzen	7	3

II. Inhaltliche Gestaltung

Fast ein Viertel aller Codings (137 von 588 Codings) bezog sich auf die „Inhaltliche Gestaltung“ von EMERGE im Sommersemester 2016 (siehe Tabelle 17). Die Studierenden diskutierten insbesondere über die Vor- und Nachbereitung der Seminartermine. Während viele Teilnehmer zur selbstständigen Nachbereitung der EMERGE-Seminare (39 Codings) die zur Verfügung gestellten Leitlinien, ein Lehrbuch oder die E-Learning-Plattform

„Amboss“ (MIAMED GmbH, Köln) genutzt haben, merkten andere Gesprächsteilnehmer an, keine konsequente Nachbereitung durchgeführt zu haben (9 Codings): „(...) *am Anfang habe ich auch mal [et]was nachgelesen aber am Ende ist es dann irgendwie wieder so unter den Tisch gefallen.*“ (FG Nr. 2, Zeile 40, weiblich). Ferner stand in der Diskussion der Gesprächsteilnehmer die Wiederholung der Patientenfälle im Fokus (50 Codings). EMERGE wurde während der letzten Seminartermine als „*sehr redundant*“ (FG2, 14, männlich) und „*langweilig*“ (FG Nr. 1, Zeile 31, weiblich) wahrgenommen. Aus diesem Grund äußerten die Studierenden den Wunsch nach einer größeren Anzahl präsentierter Patientenfälle (20 Codings).

Tabelle 17: Ergebnisse der Fokusgruppengespräche – Inhaltliche Bewertung der virtuellen Notaufnahme

Code	Codings	Interviews
II. Inhaltliche Gestaltung	137	3
II.1 Vor- und Nachbereitung der Seminare	65	3
II.1.1 Nachbereitung	60	3
II.1.1.1 Nachbereitung durchgeführt	39	3
II.1.1.1.a Nachbereitung mithilfe der bereitgestellten Leitlinien	13	3
II.1.1.1.b Nachbereitung mithilfe von Lehrbuch	9	3
II.1.1.1.c Nachbereitung mithilfe von Amboss	9	3
II.1.1.2 Keine konsequente Nachbereitung durchgeführt	9	3
II.2 Wiederholung von Fällen	50	3
II.2.1 Redundanz	45	3
II.2.1.1 Wunsch nach größerer Anzahl präsentierter Fälle	20	3
II.2.1.1.a Verärgerung über Begrenzung der präsentierten Fälle	10	3
II.2.1.2 Langweile während der letzten Seminartermine	10	3
II.3 Wünsche für die Zukunft	11	3
II.3.1 Größere Bandbreite von abgedeckten Fachgebieten	9	3
II.3 Spezifische Fälle, die in Erinnerung geblieben sind	9	2

III. Organisation des Lehrformates

Bezüglich der „Organisation des Lehrformates“ (71 Codings) (siehe Tabelle 18) diskutierten die Gesprächsteilnehmer insbesondere über die Idee, EMERGE in Kleingruppen gemeinsam zu spielen (23 Codings). Es sei „*sinnvoll, dass man (...) zusammen in einem Raum sitzt und (...) auch mal den Nachbarn irgendetwas fragen kann*“ (FG Nr. 2, Zeile 67, männlich). Eine Bearbeitung der Patientenfälle in Gruppen mit der Konsequenz, zeitweise anderen Studierenden beim Spielen zuzuschauen, ist jedoch unerwünscht. Die Gesprächsteilnehmer

schlugen stattdessen vor, in regelmäßigen Abständen eine Nachbesprechung in der Gruppe anzubieten mit der Möglichkeit „Fragen zu stellen“ und „zu diskutieren“ (FG Nr. 1, Zeile 103, weiblich). Außerdem tauschten die Studierenden ihre Meinungen bezüglich der Betreuung von EMERGE (20 Codings) aus. Sie schätzten die Möglichkeit der Diskussion von Fragen mit dem ärztlichen Betreuer als sehr hilfreich ein, übten jedoch Kritik am Betreuungsverhältnis: *"Also ich fand, muss ich ganz ehrlich sagen, das Betreuungsverhältnis relativ schlecht."* (FG Nr. 1, Zeile 21, männlich).

Tabelle 18: Ergebnisse der Fokusgruppengespräche – Organisation des Lehrformates

Code	Codings	Interviews
III. Organisation des Lehrformates	71	3
III.1 EMERGE als Gruppenarbeit	23	3
III.1.1 Kein Zuschauen beim Spielen anderer Teilnehmer gewünscht	7	3
III.1.2 Wunsch nach Nachbesprechung einiger Fälle in der Gruppe	7	2
III.1.3 Hilfreiche Beratung mit dem Sitznachbarn	7	3
III.2 Betreuung	20	3
III.2.1 Hilfreiche Diskussion von Fragen mit dem ärztlichen Betreuer	11	3
III.2.2 Schlechtes Betreuungsverhältnis	9	3
III.3 Implementierung von EMERGE im Modul 6.x	16	3
III.3.1 Repetitionscharakter von EMERGE	12	3
III.3.1.1 Guter Repetitionscharakter	10	3
III.4 Wünsche für die Zukunft	8	2
III.4.1 Oberthemen für die Seminare	8	2

IV. Emotionen

Im Rahmen der Fokusgruppendifkussionen kamen wiederholt auch Emotionen zur Sprache (55 Codings; siehe Tabelle 19). Hinsichtlich positiver Gefühle (32 Codings) berichtete ein Gesprächsteilnehmer: *„Also ich fand es auch größtenteils sehr gut, also mit sehr positiven Emotionen besetzt. Sei es irgendwie, dass man konzentriert war, dass man Spaß hatte.“* (FG3, 88). Negative Emotionen (10 Codings) entstanden überwiegend aufgrund fehlerhafter oder schwieriger technischer Aspekte der Computersimulation. Stress (13 Codings) empfanden die Studierenden hauptsächlich während der ersten EMERGE-Termine. Eine Gesprächsteilnehmerin erklärte, dass man am *„Anfang sehr aktiviert war und am Ende einfach ruhiger wurde.“* (FG Nr. 1, Zeile 29).

Tabelle 19: Ergebnisse der Fokusgruppengespräche – Erlebte Emotionen während der EMERGE-Termine

Code	Codings	Interviews
IV. Emotionen	55	3
IV.1 Positiv	32	3
IV.1.1 Spaß	15	3
IV.2 Stress	13	3
IV.2.1 Gestresst gefühlt	8	3
IV.3 Negativ	10	3
IV.3.1 Missgelaunt wegen Technik des Spiels	7	2

V. Technische Aspekte

Ein weiterer Themenschwerpunkt umfasst „Technische Aspekte“ (50 Codings) der Computersimulation (siehe Tabelle 20). In erster Linie wünschten sich die Studierenden zukünftig die Möglichkeit der selbstständigen Befundung von Untersuchungsergebnissen (11 Codings) und eine Optimierung der Handhabung (14 Codings).

Tabelle 20: Ergebnisse der Fokusgruppengespräche – Technische Aspekte

Code	Codings	Interviews
IV. Technische Aspekte	50	3
IV.1 Wünsche für die Zukunft	27	2
IV.1.1 Selbstständige Befundung von Untersuchungsergebnissen	11	2
IV.1.1.1 Selbstständige Auswertung von EKG-Aufzeichnungen	8	2
IV.2 Handhabung des Spiels	18	3
IV.2.1 Anfangs schwierige Handhabung	14	3

4 Diskussion

4.1 Die wesentlichen Ergebnisse

Im Rahmen der vorliegenden Arbeit wurde das ‚*Serious Game*‘ EMERGE als innovative Methode zur Vermittlung ärztlicher Fähigkeiten in die curriculare Lehre der UMG implementiert. Es handelt sich um eine der ersten prospektiven Interventionsstudien, welche untersucht hat, ob ein ‚*Serious Game*‘ erfolgreich zum Training von ‚*clinical reasoning*‘-Fähigkeiten in der medizinischen Lehre eingesetzt werden kann. Zudem wurde der Lernerfolg von Studierenden verglichen, die im Rahmen der curricularen Lehre entweder ein ‚*Serious Game*‘ genutzt oder die gleichen Inhalte im Rahmen eines traditionellen Lehrformats, des problemorientierten Kleingruppenunterrichts, gelernt haben.

Auf die unter 1.6 vorgestellten Studienfragen ergaben sich folgende Antworten:

- (1) Im Hinblick auf differentialdiagnostische und –therapeutische Kompetenzen erzielten Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE genutzt haben, mindestens gleichwertige Ergebnisse wie Studierende, die dieselben Inhalte im Rahmen eines problemorientierten Kleingruppenunterrichts gelernt haben.
- (2) Die in EMERGE erlernten differentialdiagnostischen und –therapeutischen Kompetenzen wiesen eine Formatspezifität auf. Hingegen war das in EMERGE erworbene Wissen weniger fallspezifisch als das im problemorientierten Kleingruppenunterricht erworbene Wissen und konnte gut auf unbekannte Fälle übertragen werden.
- (3) Im Behandlungsablauf von virtuellen Patienten ähnelten sich die Vorgehensweisen von Studierenden, die mithilfe von EMERGE oder problemorientiertem Kleingruppenunterricht gelernt haben. Beide Gruppen erhoben zuerst die Anamnese. Anschließend führten die Studierenden der EMERGE-Gruppe am häufigsten zunächst Laboranforderungen und/oder apparative Untersuchungen durch, während die Studierenden des problemorientierten Kleingruppenunterrichts die körperliche Untersuchung vorzogen.
- (4) Bei der gleichzeitigen Behandlung mehrerer virtueller Patienten empfanden Studierende, die im Rahmen der curricularen Lehre EMERGE genutzt haben, signifikant weniger Stress als Studierende, die im Rahmen der Lehre im problemorientierten Kleingruppenunterricht gelernt haben.
- (5) Studierende schätzten das ‚*Serious Games*‘ EMERGE als sehr gut geeignet zur Vermittlung ärztlicher Kompetenzen im Medizinstudium ein.

In den folgenden Kapiteln werden die Ergebnisse der hier vorgestellten Studie bezüglich der in 1.6 aufgeworfenen Studienfragen eingeordnet und diskutiert.

4.2 Effektivität von EMERGE hinsichtlich differentialdiagnostischer und –therapeutischer Kompetenzen (Studienfrage 1)

In der Literatur finden sich bereits einige Studien, die die Effektivität von ‚*Serious Games*‘ als außercurriculares Angebot in der medizinischen Lehre evaluiert haben (Youngblood et al. 2008; Diehl et al. 2015). Im Gegensatz dazu war dies die erste Forschungsarbeit, in deren Rahmen ein ‚*Serious Game*‘ in die offizielle Pflichtlehre eines medizinischen Curriculums implementiert und in welcher die Effektivität im Hinblick auf das ‚*clinical reasoning*‘-Fähigkeiten mit dem Lernen derselben Inhalte im problemorientierten Kleingruppenunterricht verglichen wurde.

Die primäre Studienfrage dieser Arbeit, ob Studierende, die im Rahmen der curricularen Lehre die Computersimulation EMERGE nutzen, bessere bzw. schlechtere Ergebnisse im Hinblick auf differentialdiagnostische und –therapeutische Kompetenzen erzielen als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts lernen, konnte angesichts der obigen Resultate (siehe Kapitel 3.3) wie folgt beantwortet werden: Die EMERGE-Teilnehmer zeigten in der ‚*Key Feature*‘-Abschlussprüfung eine signifikant bessere Leistung als die POLEMA-Teilnehmer im Hinblick auf den Gesamtpunktwert und auf die zwei Einzelfälle „Herzinsuffizienz“ und „Hyponatriämie“. Im Gegensatz zu den POLEMA-Teilnehmern waren die EMERGE-Teilnehmer allerdings während der Trainingsphase gegenüber diesen zwei Fällen exponiert worden. Hieraus lässt sich ableiten, dass eine Exposition gegenüber speziellen Lerninhalten in EMERGE grundsätzlich mit Lernerfolgen assoziiert ist. Darüber hinaus zeigte sich bezüglich derjenigen Inhalte, die zuvor sowohl in EMERGE als auch in POLEMA thematisiert worden waren (d. h. der Fälle „Sarkoidose“ und „Morbus Hodgkin“), kein signifikanter Unterschied im Hinblick auf die erreichte Gesamtpunktzahl. Es ist ein bemerkenswertes Ergebnis, dass die Studierenden der EMERGE-Gruppe den Studierenden der POLEMA-Gruppe in diesen Fällen hinsichtlich differentialdiagnostischer und –therapeutischer Kompetenzen nicht unterlegen waren, denn in EMERGE waren die Studierenden während der Interventionsphase gegenüber mindestens 23 verschiedenen Fällen exponiert worden. In POLEMA hingegen wurden innerhalb derselben Lernzeit insgesamt lediglich fünf Fälle im Detail thematisiert, womit diese zwei Prüfungsfälle bereits 40% der in POLEMA diskutierten Inhalte ausmachten.

In zwei weiteren ‚Key Feature‘-Prüfungsfällen („Fieber in Aplasie“ und „Hyperthyreose“), die zuvor nur in EMERGE thematisiert worden waren, unterschieden sich die Leistungen der Studierenden in den Gruppen EMERGE und POLEMA ebenfalls nicht signifikant voneinander. Im Fall „Fieber in Aplasie“ lag die prozentuale Leistung der EMERGE-Teilnehmer bei $68,3 \pm 23,0\%$ und die prozentuale Leistung der POLEMA-Teilnehmer bei $60,3 \pm 26,9\%$ ($p=0,096$), im Fall „Hyperthyreose“ bei $54,2 \pm 26,5\%$ bzw. $51,5 \pm 25,3\%$ ($p=0,480$). Auch wenn in diesen beiden Fällen insgesamt nur ein moderates Leistungsniveau erreicht wurde, lässt dieses Ergebnis vermuten, dass die Studierenden beider Gruppen die differentialdiagnostischen und –therapeutischen Kompetenzen bezüglich dieser Krankheitsbilder unabhängig von ihrer Teilnahme an EMERGE bzw. POLEMA erworben haben. Allerdings besteht eine Limitation dieser Studie darin, dass kein Prätest vor der Interventionsphase durchgeführt wurde, um den vorherigen Leistungsstand der Studierenden hinsichtlich differentialdiagnostischen und –therapeutischen Wissens zu evaluieren. Auch wenn die Studierenden beider Gruppen sich nicht bezüglich ihrer Klausurleistungen im Vorsemester unterschieden haben, ist es prinzipiell möglich, dass die Gruppen zu Beginn der Interventionsphase über ein unterschiedliches krankheitsspezifisches differentialdiagnostisches und –therapeutisches Wissen verfügten. Somit kann auf Grundlage der Ergebnisse dieser Studie keine sichere Aussage getroffen werden, ob entsprechende Lernerfolge tatsächlich während der Interventionsphase entstanden und somit auf die Intervention zurückzuführen sind.

Kritisch betrachtet werden sollte außerdem, dass beide Gruppen in der Abschlussprüfung insgesamt lediglich moderate Ergebnisse erzielten. Die prozentuale Leistung der POLEMA-Gruppe lag zwischen $44,9 \pm 22,0\%$ im Fall „Hyponatriämie“ und $61,8 \pm 24,8\%$ im Fall „Morbus Hodgkin“, die der EMERGE-Gruppe zwischen $54,2 \pm 26,5\%$ im Fall „Hyperthyreose“ und $68,3 \pm 23,0\%$ im Fall „Fieber in Aplasie“. Vermutlich ist der formative Charakter dieser Prüfung der Grund für die moderaten Prüfungsleistungen. Einerseits stellte die formative Prüfung für die Studierenden keinen Anlass zur gezielten Prüfungsvorbereitung dar, welche einen wahren Effekt der Intervention hätte überdecken können (Raupach et al. 2013). Andererseits waren die Studierenden beider Gruppen möglicherweise nicht motiviert, eine ihrer tatsächlichen Kompetenz entsprechende Leistung zu zeigen.

Trotz der moderaten Ergebnisse erzielte die EMERGE-Gruppe mit einer durchschnittlichen Gesamtpunktzahl von $61,8 \pm 14,5\%$ ein prinzipiell mit den Ergebnissen von anderen Studien zum Einsatz von ‚Serious Games‘ in der medizinischen Lehre vergleichbares Ergebnis. In einer

Studie von Dankbaar et al. (2017) lag beispielsweise die mittlere Gesamtpunktzahl der Interventionsgruppe nach der Nutzung eines ‚*Serious Game*‘ im Rahmen eines 10-wöchigen Praktikums in der Inneren Medizin bei $60,0 \pm 6,7\%$ in der formativen Abschlussprüfung. Ein etwas besseres Ergebnis erzielte die Interventionsgruppe mit einer durchschnittlichen Gesamtpunktzahl von 74% in der formativen Abschlussprüfung in einer Studie von Kononowicz et al. (2012).

4.2.1 Vergleich mit der bestehenden Literatur

Dass ‚*Serious Games*‘ grundsätzlich als effektives Lehrformat eingesetzt werden können und im Hinblick auf den Lernerfolg zum Teil Vorteile gegenüber traditionellen Lehrformaten zeigen, haben jüngere Meta-Analysen nachgewiesen (Cook et al. 2010; Sitzmann 2011; Connolly et al. 2012; Girard et al. 2013). Sitzmann (2011) zeigte beispielsweise in ihrer Meta-Analyse zu ‚*Serious Games*‘ (n=55), dass Studierende, die im Rahmen der Lehre ein ‚*Serious Game*‘ genutzt haben, über ein 11% höheres deklaratives Wissen, über ein 14% höheres prozedurales Wissen und über eine 20% höhere Selbstwirksamkeit verfügten als Studierende, die nicht mithilfe eines ‚*Serious Game*‘ gelernt haben. Limitierend identifizierte die Autorin allerdings Anhaltspunkte für eine Verzerrung der Ergebnisse durch Publikationsbias, welcher möglicherweise zur Überschätzung des Lernerfolgs durch ‚*Serious Games*‘ geführt hat.

Die Ergebnisse dieser Arbeit stehen auch mit konkret in der medizindidaktischen Forschung durchgeführten Studien im Einklang, die zeigen, dass ‚*Serious Games*‘ im Hinblick auf den Erwerb ärztlicher Kompetenzen, insbesondere auf den Erwerb von ‚*clinical reasoning*‘-Fähigkeiten, als mindestens gleichwertig im Vergleich zu traditionellem Unterricht einzustufen sind (Youngblood et al. 2008; Diehl et al. 2015; Dankbaar et al. 2016a; Dankbaar et al. 2017). In einer Studie von Diehl et al. (2015) wurde zum Beispiel das ‚*Serious Game*‘ ‚*InsuOnLine*‘ zum Erwerb von Wissen und Fähigkeiten zur Insulintherapie eingesetzt. An der Studie nahmen 24 Medizinstudierende aus dem 3.-6. Studienjahr und 17 Assistenzärzte der Inneren Medizin teil. Die Interventionsgruppe (n=18; 5 Assistenzärzte) zeigte mit einer Steigerung von 61% im Prätest auf 90% im Posttest eine signifikant stärkere Verbesserung ihrer Fähigkeiten als die traditionell anhand von Vorlesungen und Fallbesprechungen unterrichtete Kontrollgruppe (n=23; 12 Assistenzärzte), die ihre Leistung von 68% auf 89% verbesserte (p=0,04). Als Limitation ist in dieser Studie die unterschiedliche Anzahl von teilnehmenden Assistenzärzten pro Gruppe zu beachten, die zumindest teilweise den signifikant besseren Ausgangswert der Kontrollgruppe erklärte (61% vs. 68%; p=0,04). Im Posttest bestand zwischen den Gruppen allerdings kein Leistungsunterschied mehr, was darauf hindeutete, dass das ‚*Serious Game*‘ möglicherweise sogar effektiver war als der

traditionelle Unterricht. Zudem war die Expositionszeit gegenüber dem ‚*Serious Game*‘ und der traditionellen Lehre mit einer Dauer von drei bis vier Stunden wesentlich geringer als etwa in der hier vorliegenden Studie, in der die Expositionszeit gegenüber EMERGE bzw. POLEMA je nach Anwesenheit der Studierenden zwischen 12 und 15 Stunden betrug. Auch in weiteren Studien betrug die Expositionszeit gegenüber Lerninhalten häufig nur zwischen 30 Minuten (Carr et al. 1999) und drei bis vier Stunden (Youngblood et al. 2008; LeFlore et al. 2012). Neben einer sehr kurzen Expositionszeit ist auch eine unterschiedlich lange Expositionszeit der Interventions- und Kontrollgruppen gegenüber den jeweiligen Lehrinterventionen eine bedeutende Limitation vieler Studien (LeFlore et al. 2012; Dankbaar et al. 2016a; Dankbaar et al. 2017). In einer Studie von Dankbaar et al. (2016a) lag beispielsweise die Expositionszeit der Interventionsgruppe gegenüber dem eingesetzten ‚*Serious Game*‘ bei vier Stunden, wohingegen die Expositionszeit der Kontrollgruppe gegenüber einem Online-Lernmodul mit nur zwei Stunden signifikant geringer war ($p \leq 0,05$). Die unterschiedliche Expositionszeit gegenüber Lerninhalten könnte in entsprechenden Studien zu einer Verzerrung der Ergebnisse geführt haben. Tatsächlich lag in der vorliegenden Studie die durchschnittliche Expositionszeit gegenüber EMERGE bei ca. 13,1 Stunden (durchschnittliche Teilnahme der Studierenden an $87,1 \pm 8,4\%$ der Seminartermine) und gegenüber POLEMA bei ca. 13,9 Stunden (durchschnittliche Teilnahme der Studierenden an $92,6 \pm 6,7\%$ der Seminartermine). Somit ergab sich kein Anhalt für eine Beeinflussung der Studienergebnisse durch unterschiedlich lange Expositionszeiten, insbesondere unter Berücksichtigung der Tatsache, dass in POLEMA offiziell insgesamt 30 Minuten für die Besprechung organisatorischer Sachverhalte, beispielsweise der Gruppen- und Raumaufteilung, aufgewendet wurden.

Für den Einsatz virtueller Patienten in der medizinischen Lehre zeigte sich in einigen Studien sogar eine Überlegenheit gegenüber traditionellen Lehrformaten (Vukanovic-Criley et al. 2008; Knight et al. 2010; Kononowicz et al. 2012; LeFlore et al. 2012). In einer randomisierten Studie von LeFlore et al. (2012) erzielte beispielsweise die Interventionsgruppe ($n=47$), die ihre Fähigkeiten bezüglich Atemwegserkrankungen bei pädiatrischen Patienten mithilfe von virtuellen Patienten trainiert hatte, in der Abschlussprüfung ein signifikant besseres Ergebnis im Hinblick auf Wissen als die Kontrollgruppe ($n=46$), die dieselben Inhalte im Rahmen einer traditionellen Vorlesung gelernt hatte ($83,9 \pm 15,0\%$ vs. $75,0 \pm 12,0\%$, $p=0,004$). Allerdings handelte es sich beim Studienkollektiv nicht um Medizinstudierende, sondern um Bachelorstudierende der Krankenpflege, wodurch die Vergleichbarkeit mit den Ergebnissen der hier vorliegenden Studie eingeschränkt ist. In weiteren Studien zeigte sich eine Überlegenheit von virtuellen

Patienten gegenüber traditioneller Lehre aber auch unter Medizinstudierenden und zwar zum Beispiel im Hinblick auf Wissen und Fähigkeiten zu Erster Hilfe (Kononowicz et al. 2012) oder zu kardiologischen Erkrankungen (Vukanovic-Criley et al. 2008).

Wie jedoch Girard et al. (2013) in einer Metaanalyse feststellten, ist es auf Grundlage der bestehenden Literatur schwierig, zu allgemeinen Aussagen bezüglich der Effektivität von ‚*Serious Games*‘ zu kommen, weil die bereits veröffentlichten Studien stark hinsichtlich Forschungsfragen, eingesetzter Methoden und Studiendesign variieren. Da sich zudem auch viele ‚*Serious Games*‘ hinsichtlich Spielart und angestrebtem Effekt bzw. Lernziel voneinander unterscheiden, empfahlen die Autoren, jedes ‚*Serious Game*‘ spezifisch und individuell hinsichtlich seiner Effektivität zu evaluieren.

Zu den weiteren Limitationen bisher publizierter Studien zu diesem Thema zählen unter anderem ein kleiner Stichprobenumfang von beispielsweise nur 30 (Youngblood et al. 2008) oder 41 Probanden (Diehl et al. 2015) und der Einsatz von ‚*Serious Games*‘ im Rahmen freiwilliger außercurricularer Zusatzveranstaltungen (Vash et al. 2007; Youngblood et al. 2008; Diehl et al. 2015). Beide Limitationen sind als verzerrende Faktoren der jeweiligen Forschungsarbeiten in Betracht zu ziehen und verringern die Aussagekraft der Studienergebnisse.

Die bedeutendste Limitation eines Großteiles der Studien besteht allerdings in der Anwendung von ‚*Multiple Choice*‘-Prüfungen zur Evaluation des Lernerfolgs im Hinblick auf ‚*clinical reasoning*‘ (Cook et al. 2008). Mithilfe von ‚*Multiple Choice*‘-Prüfungen wurde zum Beispiel in Studien von Diehl et al. (2015), LeFlore et al. (2012) und Vukanovic-Criley et al. (2008) der durch ein ‚*Serious Game*‘ erzielte Lernerfolg bezüglich ärztlicher Kompetenzen untersucht. Auch in einer randomisierten Studie von Lyon et al. (1991) wurde der Lernerfolg von Medizinstudierenden im zweiten Studienjahr, die zum Erwerb von Wissen und Fähigkeiten bezüglich der Krankheitsbilder „Anämie“ und „Brustschmerz“ das computerbasierte und mit Spielelementen versehene Lehrformat „PlanAlyzer“ genutzt hatten, anhand der Beantwortung von ‚*Multiple Choice*‘-Fragen bemessen. Sowohl in der Interventionsgruppe als auch in der Kontrollgruppe, die dieselben Inhalte papierbasiert gelernt hatte, zeigten sich signifikante Lernerfolge im Prä-Postvergleich. Allerdings ergaben sich zwischen beiden Gruppen weder im Prä- noch im Posttest signifikante Leistungsunterschiede. Die Autoren selbst äußerten Bedenken darüber, dass die papierbasierte ‚*Multiple Choice*‘-Prüfung möglicherweise nicht adäquat diejenigen integrativen, höheren, kognitiven Funktionen abgebildet hat, welche die Studierenden mithilfe von „PlanAlyzer“ trainiert hatten. Denn zur Prüfung von prozeduralem Wissen, darunter zum

Beispiel Problemlösungs- und Entscheidungskompetenz im Sinne von ‚*clinical reasoning*‘, müssen Prüfungsformate gewählt werden, die die Anwendung entsprechender Fähigkeiten auch tatsächlich erfordern (siehe Kapitel 1.5.2). Besser als ‚*Multiple Choice*‘-Fragen, die eher Lernerfolge bezüglich deskriptiven Faktenwissens abbilden können, eignen sich hierzu beispielsweise ‚*Key Feature*‘-Fragen (Hrynchak et al. 2014).

4.2.2 Einsatz von ‚*Key Feature*‘-Fällen zur Prüfung differentialdiagnostischer und –therapeutischer Kompetenzen

Auch in der hier vorliegenden Studie wurde eine ‚*Key Feature*‘-Prüfung gewählt, um die Effektivität von EMERGE bzw. POLEMA im Hinblick auf differentialdiagnostische und –therapeutische Kompetenzen zu untersuchen. Da sich ‚*Key Feature*‘-Fragen klassischerweise „auf die Differentialdiagnosen, auf die diagnostischen Untersuchungen, die zur weiteren Abklärung der Diagnose nötig sind, und auf das Management und die therapeutischen Entscheidungen“ (Kopp et al. 2006) von klinischen Problemen beziehen, prüfen sie prinzipiell grundlegende Komponenten von ‚*clinical reasoning*‘. Publierte Studien, in denen ‚*Key Feature*‘-Prüfungen zur Lernerfolgsevaluation im Hinblick auf ‚*clinical reasoning*‘ angewendet wurden, wiesen eine hohe Konstruktvalidität auf und konnten erfolgreich die Auswirkungen von Lehrinterventionen auf ‚*clinical reasoning*‘-Fähigkeiten messen (Hrynchak et al. 2014). In einer ‚*Think aloud*‘-Studie von Schuwirth et al. (2001) zeigte sich, dass Studierende beim Lösen von ‚*Key Feature*‘-Problemen komplexere Gedankengänge durchliefen als beim Lösen von ‚*Multiple Choice*‘-Fragen, weshalb geschlossen wurde, dass ‚*Key Feature*‘-Fragen höhere, kognitive Leistungen erfordern als das bloße Abrufen von Faktenwissen zur Beantwortung von ‚*Multiple Choice*‘-Fragen.

Um reliable Ergebnisse bezüglich ‚*clinical reasoning*‘-Fähigkeiten zu generieren, müssen Studierende mehrere ‚*Key Feature*‘-Fälle bearbeiten (Van der Vleuten und Newble 1995; Kopp et al. 2006). Page und Bordage (1995) erzielten durch die Präsentation von 40 ‚*Key Feature*‘-Fällen in einer Prüfung eine Reliabilität von 0,8. Bei einer durchschnittlichen Bearbeitungszeit von ca. 6 Minuten pro Fall dauerte die Prüfung insgesamt ca. 4 Stunden. Ihre Erkenntnisse gewannen die Autoren aus drei Pilotstudien, die während einer sechsjährigen Forschungsarbeit von 1986-1992 durchgeführt wurden. Insgesamt wurden im Rahmen der Studie 108 ‚*Key-Feature*‘-Fälle bzw. 310 Fragen, die schließlich im Rahmen des ‚*Canadian Qualifying Examination in Medicine*‘ eingesetzt wurden, von insgesamt 3351 Studierenden pilotiert.

Im Gegensatz dazu waren der Durchführung der ‚*Key Feature*‘-Prüfung in der hier vorliegenden Studie logistische Grenzen gesetzt, die es nicht zuließen, eine vierstündige,

formative ‚*Key Feature*‘-Prüfung abzuhalten. Wir entschieden uns dazu, den Studierenden sechs ‚*Key Feature*‘-Fälle mit je vier Fragen zu präsentieren, da neuere Studien gezeigt haben, dass auch mit einer geringeren Anzahl von Fällen eine zufriedenstellende Reliabilität erreicht werden kann (Fischer et al. 2005) ohne die Studierenden mit der Länge der Prüfung zu überfordern (Kopp et al. 2006). In Übereinstimmung mit den Empfehlungen von Farmer und Page (2005) entschieden wir uns zudem für einen computerbasierten Einsatz der Prüfung, um ein ‚*backward cuing*‘, d. h. die Gabe von Lösungshinweisen in nachfolgenden Informationen bzw. Fragen (Kopp et al. 2006), zu verhindern. Das Computersystem erlaubte den Studierenden nämlich nicht, bereits getätigte Antworten nach dem Lesen von Lösungshinweisen in nachfolgenden Informationen nachträglich zu verändern. Hierdurch konnte die Authentizität der Prüfung gesteigert werden (Kopp et al. 2006), da auch voneinander abhängige Fragen gestellt werden konnten.

Schließlich wählten wir das ‚*Long Menu*‘-Antwortformat, welches Page und Bordage (1995) bei differentialdiagnostischen und –therapeutischen ‚*Key Feature*‘-Fragen als praktikable Alternative zum ‚*Write In*‘-Format empfahlen. Studienergebnisse von Rotthoff et al. (2006) zeigten, dass sich die Leistungen von Studierenden, die bei der Beantwortung derselben ‚*Key Feature*‘-Fragen entweder das ‚*Long Menu*‘-Format oder das ‚*Write In*‘-Format genutzt haben, mit einer mittleren Leistung von $73,3 \pm 16,7\%$ im ‚*Long Menu*‘-Format vs. $73,5 \pm 19,2\%$ im ‚*Write In*‘-Format nicht signifikant voneinander unterschieden ($p=0,92$). Das ‚*Long Menu*‘-Antwortformat erfordert zudem, ähnlich dem ‚*Write In*‘-Format, die Generierung spontaner Antworten (Rotthoff et al. 2006). Da dies dem ‚*clinical reasoning*‘ in der klinischen Realität ähnelt, kann durch das ‚*Long Menu*‘-Antwortformat die Authentizität und Validität einer ‚*Key Feature*‘-Prüfung gesteigert werden (Rotthoff et al. 2006).

Eine Limitation der vorliegenden Studie besteht darin, dass keine Pilotierung der ‚*Key Feature*‘-Fälle stattgefunden hat. Diese wird empfohlen, um die Inhaltsvalidität einer Prüfung zu erhöhen und die Qualität und Vollständigkeit der Synonyme und Distraktoren in der ‚*Long Menu*‘-Antwortliste zu überprüfen (Page und Bordage 1995; Kopp et al. 2006). Eine Pilotierung war aus Zeitgründen im Rahmen der vorliegenden Studie nicht möglich. An der Erstellung der ‚*Key Feature*‘-Fragen waren jedoch mehrere mit dem Prüfungsformat und der Entwicklung von ‚*Key Feature*‘-Fällen erfahrene Autoren beteiligt. Zudem waren die eingesetzten ‚*Long Menu*‘-Listen bereits im Rahmen mehrerer Forschungsarbeiten pilotiert und erprobt worden (Raupach et al. 2016; Ludwig et al. 2018). Durch die Erstellung einer nationalen Datenbank mit pilotierten ‚*Key Feature*‘-Fällen und Antwortlisten ließe sich der Einsatz des Prüfungsformates ökonomischer gestalten (Kopp et al. 2006).

Eine weitere Limitation stellt die Auswahl der sechs ‚Key Feature‘-Fälle dar, die im Rahmen der Abschlussprüfung präsentiert wurden: Jeweils zwei Fälle bezogen sich auf die inhaltlichen Schwerpunkte der EMERGE-Gruppe A bzw. B, waren aber in POLEMA zuvor nicht diskutiert worden. Zwei weitere Fälle stammten aus der inhaltlichen Schnittmenge zwischen EMERGE und POLEMA. In einer zukünftigen Studie sollten zusätzlich ‚Key Feature‘-Fälle präsentiert werden, die zuvor nur in POLEMA, nicht aber in EMERGE, diskutiert worden waren, um die Leistung der EMERGE-Gruppe bezüglich inhaltlich unbekannter Fälle im Vergleich zur POLEMA-Gruppe untersuchen zu können.

4.2.3 ‚Blended learning‘

Die beiden Lehrformate EMERGE und POLEMA waren im Rahmen eines ‚Blended learning‘-Ansatzes in das didaktische Gesamtkonzept des neuen Repetitionsmoduls 6.x eingebettet. ‚Blended learning‘ ermöglicht Studierenden ein hohes Maß an Unabhängigkeit im Lernprozess und regt sie zu kritischem Diskurs und reflexivem Denken an (Garrison und Kanuka 2004). Da auf diese Weise komplexe Gedankenprozesse und kritisches Denken gefördert werden (Garrison und Kanuka 2004), bot ‚Blended learning‘ optimale Bedingungen für die Vermittlung und das Training von ‚clinical reasoning‘ im Rahmen des Repetitionsmoduls 6.x. Allerdings ging mit diesem Ansatz einher, dass thematisierte Inhalte nicht nur in EMERGE bzw. POLEMA, sondern auch in Vorlesungen und Seminaren, die während des Moduls stattfanden, angesprochen wurden. Eine unterschiedlich häufige Anwesenheit der Studierenden bei entsprechenden Veranstaltungen mit einer entsprechend unterschiedlich starken Exposition gegenüber ebendiesen Inhalten könnte die Ergebnisse dieser Studie beeinflusst haben. Die Anwesenheit bei Vorlesungen und Seminaren wurde im Rahmen dieser Studie aber nicht erhoben, sodass diese potentielle Verzerrung nicht quantifizierbar ist. Jedoch ergab sich weder aus dem vorhandenen quantitativen Datenmaterial noch auch aus den Fokusgruppengesprächen ein Anhalt dafür, dass es aufgrund des ‚Blended learning‘-Ansatzes zur Benachteiligung einer Gruppe gekommen sein könnte.

Einen weiteren potentiellen Störfaktor stellt das Selbststudium der Studierenden dar. Im Rahmen dieser Studie wurden keine Daten dazu gesammelt, wie viel Zeit für selbstständiges Lernen außerhalb der Lehrveranstaltungen aufgewendet wurde, sodass dessen Auswirkung auf die Ergebnisse unklar blieb. Einzelne Studienergebnisse gaben Hinweise darauf, dass ‚Serious Games‘ in einem stärkeren Ausmaß zum Selbststudium motivierten als traditionelle Lehrformate (Dankbaar et al. 2016a; Dankbaar et al. 2017). In einer Studie von Dankbaar et al. (2016b), beispielsweise, wendete die Interventionsgruppe (n=107) während der Vorbereitung auf einen Erste Hilfe-Kurs mithilfe des ‚Serious Game‘ ‚abcdeSIM‘ zweieinhalb

Stunden mehr für selbstständiges Lernen auf als die Kontrollgruppe (n=52), die zur Vorbereitung ein Skript gelesen hat ($p=0,007$). Allerdings ist die Aufbereitung der Lehrinhalte als reiner Lesestoff keineswegs mit dem in dieser Studie als Vergleich zu EMERGE herangezogenen Lehrformat POLEMA vergleichbar. Da die POLEMA-Teilnehmer explizit dazu aufgefordert wurden, die Zeit zwischen zwei Seminarterminen zur Aneignung von detailliertem Wissen bezüglich thematisierter Inhalte zu nutzen, kann angenommen werden, dass sie mindestens ebenso viel, wenn nicht sogar mehr Zeit mit selbstorganisiertem Lernen verbracht haben als die EMERGE-Teilnehmer. Zudem wiesen die Fokusgruppengespräche auf eine heterogene Intensität des Selbststudiums unter den EMERGE-Teilnehmern hin: Zwar haben einige EMERGE-Teilnehmer ihre Seminare beispielsweise mithilfe von Leitlinien oder Lehrbüchern selbstständig nachgearbeitet (39 Codings), jedoch führten andere Studierende keinerlei Nachbereitung durch (9 Codings). Wir empfehlen für zukünftige Studien, Daten bezüglich der Intensivität und Dauer des Selbststudiums in beiden Gruppen zu erheben, um einen möglichen Einfluss auf die Studienergebnisse evaluieren zu können. Zudem könnte in einem randomisierten Design ein strukturiert angeleitetes Selbststudium als möglicherweise effektivitätssteigernde Intervention untersucht werden.

4.3 Format- und/oder Fallspezifität der erlernten Kompetenzen (Studienfrage 2)

Die Ergebnisse der EMERGE-Abschlussprüfung zeigten, dass die während der Trainingsphase in EMERGE erlernten Kompetenzen eine hohe Formatspezifität aufwiesen. Ferner wurde erkennbar, dass die in EMERGE erlernten Kompetenzen wesentlich weniger fallspezifisch und damit in höherem Ausmaß auf neue Fälle übertragbar waren als die in POLEMA erlernten Kompetenzen.

4.3.1 Formatspezifität

Die EMERGE-Teilnehmer übertrafen in der EMERGE-Abschlussprüfung deutlich die Leistungen der POLEMA-Teilnehmer. Erstens erreichten sie in allen vier Prüfungsfällen, die zuvor in keinem der beiden Lehrformate diskutiert worden waren, eine signifikant höhere Gesamtpunktzahl. Zweitens waren ihre Leistungen im Hinblick auf alle vier Fälle in mehreren Kategorien, u. a. „Korrekte Diagnose“ und „Korrekte Therapie“, signifikant besser als die Leistungen der POLEMA-Teilnehmer. Der signifikante Vorteil der Computersimulation gegenüber dem problemorientierten Lernen könnte allerdings auf

einem reinen Trainingseffekt beruht haben. Möglicherweise hatten die POLEMA-Teilnehmer Probleme mit der Steuerung der Computersoftware. Dieser Verdacht ergab sich aus den Fokusgruppengesprächen, in denen einige EMERGE-Teilnehmer anmerkten, dass die Handhabung des Spiels anfangs schwierig (14 Codings) aber dennoch intuitiv und schnell zu erlernen (4 Codings) gewesen sei. In Folgestudien sollte eine Trainingseinheit angeboten werden, um die POLEMA-Teilnehmer mit der Software vertraut zu machen. Allerdings erforderte beispielsweise die Auswahl der richtigen Diagnose per Mausclick aus einer alphabetischen Liste von ca. 130 möglichen Diagnosen kein spezifisches Training. Trotz dessen stellten die EMERGE-Teilnehmer in drei von vier Prüfungsfällen („NSTEMI“, „Pankreatitis“, „Asthmaanfall“) signifikant häufiger die richtige Diagnose als die POLEMA-Teilnehmer.

4.3.2 Fallspezifität

Die EMERGE-Teilnehmer waren sehr gut dazu in der Lage, ihre während der Trainingsphase erworbenen differentialdiagnostischen, -therapeutischen und prozeduralen Fähigkeiten auf die Bearbeitung unbekannter Patientenfälle in der EMERGE-Abschlussprüfung zu übertragen. Im Gegensatz dazu waren die im problemorientierten Kleingruppenunterricht erlernten Fähigkeiten sehr viel fallspezifischer und konnten von den POLEMA-Teilnehmern weniger gut auf die Behandlung virtueller Patienten mit unbekanntem Krankheitsbildern in der EMERGE-Abschlussprüfung angewendet werden.

Zwischen den drei EMERGE-Gruppen A, B und AB zeigten sich unter insgesamt 31 Items, die im Rahmen der Auswertung der EMERGE-Abschlussprüfung untersucht wurden, lediglich zwei signifikante Leistungsunterschiede. Studierenden aller Gruppen wurden ähnliche Konsultationsanlässe bzw. Symptomkomplexe präsentiert. Obwohl die zugrundeliegenden Erkrankungen derjenigen virtuellen Patienten, die den Studierenden der EMERGE-Gruppe A präsentiert worden waren, hauptsächlich aus den Fachbereichen Kardiologie/Pulmologie und Nephrologie/Rheumatologie stammten (siehe Tabelle 4), waren ihre Leistungen im Hinblick auf die gastroenterologisch-endokrinologischen Prüfungsfälle „Pankreatitis“ und „Untere gastrointestinale Blutung“ nicht schlechter als die Leistungen der EMERGE-Gruppen B. Ebenso erreichten die Studierenden der EMERGE-Gruppe B, die während der Trainingsphase hauptsächlich Patienten mit Erkrankungen aus den Fachbereichen Gastroendokrinologie und Hämatologie/Onkologie bearbeitet hatten, in den kardiologisch/pulmologischen Prüfungsfällen „Asthmaanfall“ und „NSTEMI“ ähnliche Punktzahlen wie die Studierenden der EMERGE-Gruppe A. Dies zeigte, dass alle EMERGE-Teilnehmer ‚*clinical reasoning*‘-Fähigkeiten zum erfolgreichen Management von

Symptomkomplexen wie beispielsweise „Fieber“ und „Bauchschmerz“ erlernt haben, und zwar unabhängig von der zugrundeliegenden Erkrankung.

Ein bemerkenswertes Ergebnis ist ferner, dass die EMERGE-Gruppe AB im Vergleich zu den Gruppen A und B in allen Fällen gleich gut abgeschnitten hat. Einerseits war diese Gruppe gegenüber Fällen aus allen vier Fachbereichen exponiert worden. Andererseits war das Ausmaß dieser Exposition aber wesentlich geringer als das jeweilige Ausmaß der Exposition in den Gruppen A und B. Die Ergebnisse geben Hinweise darauf, dass die Studierenden trotzdem Fähigkeiten und Wissen in derselben Tiefe wie ihre Kommilitonen erlangt haben. Da die EMERGE-Teilnehmer in den Fokusgruppengesprächen zudem den Wunsch nach einer größeren Anzahl präsentierter Patientenfälle äußerten (20 Codings), wird für zukünftige Studien vorgeschlagen, die Redundanz zu minimieren, weil diese den Lernerfolg nicht erhöht. Wie stark die Exposition gegenüber Fällen tatsächlich sein muss, um optimale Lernerfolge zu erzielen, sollte in zukünftigen Studien untersucht werden.

Auffällig ist allerdings, dass zwischen den Leistungen bei den vier EMERGE-Prüfungsfällen deutliche Unterschiede bestanden. Insbesondere beim Fall „Untere gastrointestinale Blutung“ erreichten die EMERGE-Teilnehmer ausgesprochen wenige Punkte bezüglich der Aspekte „Korrekte Diagnose“ (58,6%; n=41) und „Korrekte Therapie“ ($14,3 \pm 35,2\%$; n=70). Dies erstaunte, da gastroenterologisch-endokrinologischen Inhalte in den EMERGE-Gruppen B und AB etwa die Hälfte bzw. ein Viertel der während der Trainingsphase präsentierten Inhalte ausmachten. Für zukünftige Studien leitet sich die Frage ab, inwieweit der spezifische, inhaltsbezogene Lernerfolg tatsächlich von der Exposition gegenüber entsprechenden Lerninhalten während der Trainingsphase abhängt.

Gegen eine Fallspezifität der in EMERGE erlernten kognitiven Kompetenzen sprachen auch die Ergebnisse der ‚Key Feature‘-Abschlussprüfung. In dieser unterschieden sich die drei EMERGE-Gruppen A, B und AB weder signifikant hinsichtlich der erreichten Gesamtpunktzahl noch hinsichtlich ihrer Leistungen in den sechs Einzelfällen.

4.4 Prozedurales Vorgehen (Studienfrage 3)

Im Hinblick auf das strukturelle Vorgehen im Behandlungsablauf virtueller Patienten konnte zunächst festgestellt werden, dass die EMERGE-Teilnehmer alle vier EMERGE-Prüfungsfälle in einem höheren Prozentsatz abgeschlossen, d. h. verlegt oder entlassen, haben als die POLEMA-Teilnehmer. Diese Ergebnisse gaben einen weiteren Hinweis darauf, dass die POLEMA-Teilnehmer möglicherweise nicht im gleichen Ausmaß mit der Handhabung der Software vertraut waren wie die EMERGE-Teilnehmer. Um den

diesbezüglichen Vorteil der EMERGE-Gruppe möglichst gering zu halten, wurden in die detaillierte Auswertung des prozeduralen Vorgehens in der EMERGE-Abschlussprüfung nur diejenigen Studierenden beider Gruppen einbezogen, die den jeweiligen Fall abgeschlossen haben.

In dieser Auswertung zeigte sich, dass in beiden Gruppen bei der Behandlung aller vier Patienten am häufigsten mit der Anamnese begonnen wurde. Anschließend bevorzugte die POLEMA-Gruppe es, zunächst eine körperliche Untersuchung gefolgt von Labor/apparativen Untersuchungen durchzuführen. Die EMERGE-Gruppe hingegen forderte am häufigsten zunächst Labor/apparative Untersuchungen an und nahm erst anschließend die körperliche Untersuchung vor.

Diese Ergebnisse können nicht mit bestehenden Studien verglichen werden, da es bislang es keine publizierte Studie gibt, in welcher Log-Dateien eines ‚*Serious Games*‘ zur Untersuchung des Behandlungsablaufs virtueller Patienten analysiert wurden. Somit kommt diesen Ergebnissen eher ein hypothesengenerierender Charakter zu.

Die Anamnese stellte bei Studierenden beider Gruppen den ersten Schritt im Behandlungsablauf dar. Dass die Anamnese unter den Studierenden einen hohen Stellenwert hat, ist gut und wichtig, da eine ausführliche Anamnese wesentlich zur Diagnosefindung beiträgt und Schätzungen zufolge in bis zu 90% die entscheidende diagnostische Methode ist (Gross und Löffler 1997).

Auffällig war, dass die EMERGE-Teilnehmer im Anschluss an die Anamnese die Anforderung von Labor/apparativen Untersuchungen priorisierten. Dieses Vorgehen muss vor dem Hintergrund, dass in mehreren Studien eine übermäßige und zum Teil unnötige Anforderung von Laborwerten und apparativen Untersuchungen durch Mediziner festgestellt werden konnte (Winkens und Dinant 2002; Miglioretti und Smith-Bindman 2011), kritisch betrachtet werden. Miyakis et al. (2006) untersuchten beispielsweise retrospektiv 426 Patientenakten auf den Nutzen von 25 spezifischen Laborwerten (u. a. hämatologische Laborwerte, arterielle Blutgasanalyse), die innerhalb eines halben Jahres angefordert wurden. Pro Patient wurden im Schnitt 2,96 Laboranforderungen pro Tag getätigt. Hiervon waren insgesamt 67,9% nicht relevant für das weitere Patientenmanagement. Die Autoren schlussfolgerten, dass 2,01 Laboranforderungen/Tag/Patient vermeidbar gewesen wären. Die Gründe für die übermäßige Anforderung von Labor/apparativen Untersuchungen durch Ärzte waren beispielsweise Unsicherheit, mangelnde Erfahrung und Ahnungslosigkeit über die entstehenden Kosten (Young 1988; Miyakis et al. 2006).

In einer Studie von Harendza et al. (2013) wurde das Anforderungsverhalten von Medizinstudierenden hinsichtlich Laborwerten und bildgebenden Verfahren zwischen Studierenden der Medizinfakultäten in Utrecht (Niederlande) und Hamburg verglichen. In Utrecht fand die Lehre im Rahmen eines ‚*vertically integrated curriculum*‘ statt, in dem von Beginn an theoretisches Wissen mit praktisch-klinischen Erfahrungen verzahnt wurde. Im Gegensatz dazu herrschte in Hamburg ein Regelstudiengang mit voneinander getrenntem vorklinischem und klinischem Abschnitt („*non-vertically integrated curriculum*“) vor (siehe Kapitel 1.1). Jeweils 30 kurz vor der Approbation stehende Studierende beider Fakultäten nahmen freiwillig an einer Prüfung teil, in welcher fünf Patienten mit verschiedenen Krankheitsbildern (Zöliakie, Wegener-Granulomatose, perforierte Sigmadivertikulitis, Myasthenia Gravis, Varizella-Zoster-Virus-Infektion) behandelt werden mussten. Im Behandlungsablauf dieser Patienten zeigte sich, dass die deutschen Medizinstudierenden insgesamt eine signifikant höhere Anzahl von Laboruntersuchungen (466 vs. 283; $p < 0,01$) und Bildgebungen (156 vs. 97; $p < 0,01$) anforderten als die niederländischen Medizinstudierenden. Kein signifikanter Unterschied bezüglich der angeforderten Anzahl von Bildgebungen ergab sich lediglich im Fall „perforierte Sigmadivertikulitis“, dem einzigen Fall, in dem ein bildgebendes Verfahren auch tatsächlich zur Diagnosefindung und Therapieplanung indiziert war. Der deutlichste Unterschied zeigte sich hingegen für den Fall eines 5-jährigen Kindes mit der Diagnose „Zöliakie“, bei dem zwei Drittel der deutschen Medizinstudierenden ohne Indikation und Notwendigkeit für den Behandlungsverlauf eine computertomographische Bildgebung anfertigen wollten (20 vs. 0; $p < 0,001$). Dies ist besorgniserregend, da eine jüngere Studie von Pearce et al. (2012) ergab, dass Kinder mit einer kumulativen Strahlendosis von 50-60 mGy aufgrund von Computertomographien in der Folge mit einer höheren Wahrscheinlichkeit an den Krankheiten „Leukämie“ und „Hirnkrebs“ erkrankten. Aus diesem Grund sollten Bildgebungen bei jungen Patienten zurückhaltend eingesetzt werden. Die Ergebnisse der Studie von Harendza et al. (2013) legen nahe, dass Studierende, die im Rahmen eines ‚*virtually integrated curriculum*‘ schon früh im Studienverlauf mit klinischen Problemen konfrontiert werden und in diesem Rahmen über viele Jahre ihre ‚*clinical reasoning*‘-Fähigkeiten trainieren können, fokussierte und verantwortungsbewusstere Anforderungen bezüglich Laborwerten und bildgebenden Verfahren tätigen als Studierende eines Regelstudiengangs.

Im Hinblick auf zukünftige Studien empfehlen wir, im Rahmen der Auswertung von Log-Dateien eines ‚*Serious Games*‘ wie EMERGE neben der Reihenfolge der getätigten Anforderungen und Untersuchungen auch die Anzahl und Notwendigkeit derselben zu untersuchen. Zudem könnten im Rahmen von Folgestudien die Kosten für

Labor/apparative Untersuchungen in EMERGE angezeigt werden. Hierdurch könnte der weiterführenden Studienfrage nachgegangen werden, ob sich die Kenntnis der Kosten bestimmter Laboranforderungen und apparativer Untersuchungen tatsächlich, wie in der Literatur vermutet wird (Young 1988), auf den Behandlungsverlauf von Patienten durch Medizinstudierende bzw. Ärzte auswirkt. Zudem könnte der Ressourceneinsatz entweder in Nachbesprechungen von EMERGE-Terminen thematisiert oder Studierende vor der Freigabe von Tests im Spiel dazu aufgefordert werden, die Anforderung vor dem Hintergrund der Vortestwahrscheinlichkeit zu begründen und anzugeben, welchen Einfluss ein spezifisches Ergebnis auf ihre Beurteilung der klinischen Situation haben würde.

4.4.1 Priorisierung von Labor/apparativen Untersuchungen im Fall „NSTEMI“

Bei der Behandlung des Falls „NSTEMI“ forderten 31,1% der EMERGE-Teilnehmer zuerst, d. h. sogar vor Durchführung der Anamnese, Labor/apparative Untersuchungen an. Auch in der POLEMA-Gruppe zeigte sich (abweichend vom Gesamtbild über alle Fälle hinweg), dass die Studierenden in diesem spezifischen Fall Labor/apparative Untersuchungen priorisierten. Zwar erhoben die Studierenden dieser Gruppe weiterhin am häufigsten zunächst die Anamnese, forderten anschließend aber häufiger zunächst Labor/apparative Untersuchungen als eine körperliche Untersuchung an. Diese Ergebnisse legen die Vermutung nahe, dass die Studierenden sowohl in EMERGE als auch in POLEMA richtigerweise gelernt haben, dass es bei „Brustschmerz“ und klinischen Anzeichen für einen „NSTEMI“ gilt, möglichst zeitnah ein EKG anzulegen und die notwendigen Laborparameter, insbesondere das kardiospezifische Isoenzym der Kreatinkinase (CK-MB) und die Troponine T und I, anzufordern (Roffi et al. 2016). In den Fokusgruppengesprächen merkten die EMERGE-Teilnehmer außerdem an, dass ihnen im Laufe der Trainingsphase bewusstgeworden war, dass es eine gewisse Zeit dauerte, bis die Ergebnisse von angeforderten Laborwerten oder apparativen Untersuchungen verfügbar waren. Die jeweiligen Latenzzeiten hielten sie für sehr realistisch im Hinblick auf den Krankenhausalltag. Sie vermerkten als Lernerfolg die Aneignung von Behandlungsalgorithmen (13 Codings), äußerten jedoch auch, gelernt zu haben, dass es in einigen Fällen sinnvoll und notwendig sein kann, von Algorithmen abzuweichen, um beispielsweise zeiteffizienter zu arbeiten.

4.5 Stressempfinden bei der Behandlung mehrerer virtueller Patienten (Studienfrage 4)

Die Ergebnisse dieser Studie zeigten erstmals, dass Studierende, die mithilfe eines ‚*Serious Game*‘ gelernt haben, bei der gleichzeitigen Behandlung mehrerer virtueller Patienten weniger Stress empfanden als Studierende, die den gleichen Inhalt im Rahmen eines problemorientierten Kleingruppenunterrichts diskutiert haben. In den Fragebögen gaben die POLEMA-Teilnehmer diesbezüglich ein signifikant höheres Stresslevel an als die EMERGE-Teilnehmer ($3,3 \pm 1,6$ vs. $4,3 \pm 1,4$; $p=0,001$), auch wenn die Mittelwerte beider Gruppen auf ein insgesamt nur mäßiges Stresslevel hinwiesen.

Einen großen Unterschied zwischen den beiden Lehrformaten stellt die Tatsache dar, dass in POLEMA zu keinem Zeitpunkt mehrere Patienten gleichzeitig behandelt werden müssen. Dies bedeutet, dass Schlüsselaspekte von ‚*clinical reasoning*‘ mithilfe von POLEMA nicht (hinreichend) trainiert werden können, beispielsweise müssen keine Patienten triagiert und keine Dringlichkeiten von Symptomkomplexen eingeschätzt werden. Gerade die Notwendigkeit, bei der gleichzeitigen Behandlung mehrerer virtueller Patienten Prioritäten zu setzen (27 Codings) und Dringlichkeiten einzuschätzen (12 Codings), wurde jedoch von den EMERGE-Teilnehmern als bedeutendster Vorteil von EMERGE gegenüber traditionellen Lehrformaten angesehen. Kongruent stimmten sie auch den Fragebogenaussagen zu, durch das Spielen von EMERGE im Anschluss an die Trainingsphase tatsächlich die Dringlichkeit klinischer Situationen ($2,4 \pm 1,2$) und die Wertigkeit verschiedener diagnostischer Schritte ($2,5 \pm 1,2$) besser einschätzen zu können. Es ist anzunehmen, dass die Lernerfolge der EMERGE-Gruppe hinsichtlich des prozeduralen Vorgehens, beispielsweise der Aneignung von Behandlungsalgorithmen (siehe Kapitel 4.4) und des Triagierens, zu einem geringeren Stressempfinden während der EMERGE-Abschlussprüfung geführt haben. Entsprechend äußerten die Studierenden in den Fokusgruppengesprächen, dass sie sich insbesondere während der ersten EMERGE-Seminartermine gestresst gefühlt haben, dieser Stress im Laufe der Trainingsphase aber durch die Entwicklung von Routinen bei der gleichzeitigen Behandlung mehrerer Patienten geringer wurde. Insofern kann der in der Befragung beobachtete signifikante Unterschied zwischen der POLEMA und der EMERGE-Gruppe lediglich ein Produkt der mangelnden Spielerfahrung in der POLEMA-Gruppe gewesen sein.

Insgesamt gaben die Ergebnisse dieser Studie Hinweise darauf, dass EMERGE sich günstig auf das Stressempfinden von Studierenden bei der gleichzeitigen Behandlung mehrerer Patienten auswirkt. Auch im Hinblick auf diese Studienfrage wäre bei weiteren Studien ein

Prätest, wie er bereits in Kapitel 4.2 vorgeschlagen wurde, sinnvoll, um das Stresslevel der Studierenden bei der gleichzeitigen Behandlung mehrerer Patienten vor der Intervention mit ihrem diesbezüglichen Stresslevel nach der Intervention vergleichen zu können. Limitierend handelte es sich im Rahmen dieser Studie um subjektive Selbsteinschätzungen der Teilnehmer. In Folgestudien sollte das Stressempfinden objektiviert werden, beispielsweise durch die Messung der Herzfrequenz oder der Ausschüttung von Stresshormonen. Zudem gaben Studien Hinweise darauf, dass Studierende bei der Interaktion mit einer Computersoftware generell wesentlich weniger Stress empfinden als bei der Interaktion mit echten Patienten (Bryce et al. 1998). In zukünftigen Studien sollte deshalb auch der Frage nachgegangen werden, wie sich das Training der gleichzeitigen Behandlung mehrerer virtueller Patienten in EMERGE auf das Stressempfinden von Studierenden bei der gleichzeitigen Behandlung mehrerer realer Patienten auswirkt.

In einer Studie von Dankbaar et al. (2017) wurde beispielsweise untersucht, wieviel Stress Medizinstudierende im vierten Studienjahr, die zuvor mithilfe des ‚*Serious Game*‘ ‚Air-Medic Sky-1“ die simultane Behandlung mehrerer Patienten trainiert hatten, bei der Behandlung von Patienten im Rahmen eines 10-wöchigen Praktikums in der Inneren Medizin empfanden. Die Studienteilnehmer (n=32) füllten wöchentlich Fragebögen basierend auf der ‚*Perceived Stress Scale*‘ nach Cohen et al. (1983) aus, in denen sie von ihrem subjektiven Stressempfinden berichteten. Die Studie ergab, dass das Stresslevel der Studierenden insgesamt lediglich moderat und damit vergleichbar mit den Ergebnissen der hier vorliegenden Studie war. Zudem ergab sich kein signifikanter Unterschied hinsichtlich des Stressempfindens zwischen der Interventionsgruppe und einer historischen Kontrollgruppe (n=37), die zuvor nicht mithilfe des ‚*Serious Game*‘ gelernt hatte. Die Autoren erklärten das insgesamt niedrige Stresslevel der Teilnehmer dadurch, dass es sich um Studierende handelte, welche die Last der Verantwortung und den Stress im klinischen Alltag noch gar nicht erlebt haben. Folglich schlugen sie vor, eine Folgestudie an einer Kohorte von Berufsanfängern durchzuführen, die die volle Verantwortung für ihr Handeln trägt und somit potentiell mehr Stress bei der Patientenbehandlung empfindet. Eine qualitative Studie von Brennan et al. (2010) unter Assistenzärzten im ersten Berufsjahr (n=31) ergab nämlich beispielsweise, dass der Übergang vom Medizinstudierenden zum Arzt häufig als extrem schwierig und stressig empfunden wurde. Die Probanden äußerten, erst in den ersten Wochen ihrer Tätigkeit als Arzt realisiert zu haben, dass sie fortan die letztendliche Verantwortung für die Behandlung und das Leben ihrer Patienten trugen, was vielen Berufsanfängern Angst machte. Ferner betonten sie, dass eine gute Vorbereitung auf die ärztliche Tätigkeit während des Medizinstudiums den Übergang vom Studium in die Praxis erleichterte. Hierzu zählten sie

beispielsweise Lehrformate, in denen sie schon als Studierende Verantwortung für Patienten übernehmen und selbstständig Priorisierungen vornehmen mussten. Da EMERGE auf Basis der Ergebnisse dieser Studie grundsätzlich als gute Vorbereitung auf den Berufsalltag erscheint, kann vermutet werden, dass sich dieses ‚*Serious Game*‘ günstig auf das Stressempfinden von Berufsanfängern auswirkt. Ob Assistenzärzte, die während ihres Studiums mit EMERGE gelernt haben, tatsächlich weniger Stress bei der Patientenbehandlung empfinden als Assistenzärzten, die ihre ‚*clinical reasoning*‘-Fähigkeiten zuvor nicht mithilfe eines ‚*Serious Game*‘ trainiert haben, sollte in Folgestudien untersucht werden.

4.6 EMERGE als geeignetes Lehrformat für den Erwerb ärztlicher Kompetenzen (Studienfrage 5)

Laut den Ergebnissen der Fragebogen-Erhebung in dieser Studie halten Studierende EMERGE grundsätzlich für ein geeignetes Lehrformat zum Erwerb ärztlicher Kompetenzen ($1,7 \pm 0,8$). In den Fokusgruppengesprächen bekräftigten die Studierenden, in Übereinstimmung mit den Studienergebnissen anderer Autoren (Bryce et al. 1998; Wahlgren et al. 2006; Diehl et al. 2015), zudem ihre Akzeptanz gegenüber ‚*Serious Games*‘ und zählten viele Vorteile von EMERGE gegenüber traditionellen Lehrformaten auf (201 Codings). Die Studierenden stellten insbesondere heraus, dass EMERGE selbständiges Arbeiten fordert und fördert (49 Codings). Dieser Aspekt wurde auch in einer semiquantitativen Fragebogenstudie (n=31) von Wahlgren et al. (2006) deutlich, deren Teilnehmer (Medizinstudierende, 7. Semester) zuvor im Rahmen eines 17-wöchigen Dermatologie- und Venerologie-Kurs zusätzlich zur traditionellen Lehre (u. a. Vorlesungen, Seminare, Unterricht am Krankenbett) die interaktive Simulationssoftware „NUDOV“ genutzt hatten. Im Gegensatz dazu hatte die Kontrollgruppe (n=85), die nicht in die Fragebogenstudie einbezogen wurde, lediglich traditionelle Lehre erhalten. Insgesamt urteilten 90% der Studierenden, dass „NUDOV“ das Lernen erleichterte und sie ihr Wissen mithilfe von „NUDOV“ schneller erweitern konnten als mithilfe von traditionellem Unterricht (84%). Darüber hinaus lobten 65% der Studierenden die Selbstständigkeit im Lernprozess und die Möglichkeit, im eigenen Tempo zu lernen, zu wiederholen und zu reflektieren.

Weiterhin hielten die Studierenden EMERGE für ein realistisches und praxisnahes Szenario (25 Codings) und für eine gute Vorbereitung auf den späteren Berufsalltag (10 Codings). Übereinstimmend ergab eine Studie von Lyon et al. (1991), dass das ‚*Serious Game*‘

„PlanAlyzer“ Studierende nach eigenen Angaben gut auf studentische Praktika und ihre zukünftige ärztliche Tätigkeit, insbesondere im Hinblick auf die Themen „Brustschmerz“ und „Anämie“, vorbereitet hat. Die Studierenden schätzten das ‚*Serious Game*‘ diesbezüglich als effektiver als das Lernen mit textbasierten Fällen ein.

Nicht zuletzt brachten die Studierenden in den Fokusgruppengesprächen im Hinblick auf die Nutzung von EMERGE positive Emotionen (32 Codings) und Spaß (15 Codings) zum Ausdruck. Vergleichbare Aussagen konnten auch in anderen Studien erhoben werden. In diesen zeigte sich, dass die Nutzung eines ‚*Serious Game*‘ in der medizinischen Lehre als angenehm empfunden wurde, den Studierenden Spaß machte und eine Motivation zum Lernen darstellte (Bryce et al. 1998; Diehl et al. 2015; Dankbaar et al. 2016b; Dankbaar et al. 2017).

Forschungsarbeiten zeigen, dass, neben kognitiven Prozessen, auch „emotional-motivationale Begleitprozesse“ entscheidenden Einfluss auf Lernprozesse haben (Krapp 2005). Emotionen haben u. a. Einwirkung auf die Aufmerksamkeit, auf die Auswahl von Lernstrategien und auf die Lernanstrengungen und stellen damit zentrale Prädiktoren für die Lernleistung dar (Frenzel et al. 2009; Pekrun 2014). Laut Pekrun (2014) haben positive Emotionen (z. B. Freude, Stolz) aktivierende Effekte auf die Lernmotivation, das Interesse und die Lernleistung. Im Gegensatz dazu wirken negative Emotionen (z. B. Langeweile) häufig deaktivierend (Pekrun 2014).

Das enorme Motivationspotenzial von ‚*Serious Games*‘ bezüglich des Erreichens von Lernzielen, beispielsweise ärztlicher Kompetenzen, wird in der Spieltheorie als ihr primärer Nutzen postuliert (Garris et al. 2002). Nach der Selbstbestimmungstheorie von Ryan und Deci (2000) wird grundsätzlich zwischen intrinsischer Motivation (d. h. eine Handlung geschieht aus Spaß, Zufriedenheit und der natürlichen Tendenz, das eigene Wissen und die eigenen Fähigkeiten zu erweitern) und extrinsischer Motivation (d. h. eine Handlung geschieht aus instrumenteller Absicht) unterschieden. Der Theorie liegt zugrunde, dass der Mensch in seinen Handlungen die folgenden drei Bedürfnisse befriedigen möchte: Kompetenz, Autonomie/Selbstbestimmung und soziale Eingebundenheit. Insbesondere Autonomie und Selbstbestimmung determinieren das Ausmaß der intrinsischen Motivation (Ryan und Deci 2000). Studien zeigten, dass Studierende, die in ihrem Lernprozess viel Autonomie und Selbstbestimmung erfuhren, zufriedener waren (Ross et al. 1989), höhere Level intrinsischer Motivation erreichten, sich tiefergehend mit Lehrinhalten beschäftigten und sich schließlich kompetenter fühlten als Studierende, denen weniger Autonomie gewährt wurde (Cordova und Lepper 1996). ‚*Serious Games*‘ fordern und fördern Autonomie, steigern

durch die im Interaktionsprozess mit dem Spiel generierten Emotionen wie Freude, Spaß und Ehrgeiz die intrinsische Motivation des Lernenden und spornen ihn zum Kompetenzerwerb und dem Erreichen von Lernzielen an (Garris et al. 2002).

4.6.1 Fokusgruppengespräche zur Evaluation von Meinungen bezüglich des Einsatzes von EMERGE in der Lehre

Wir entschieden uns, Fokusgruppengespräche durchzuführen, um die Studierenden nach ihren Erfahrungen bezüglich des Einsatzes von EMERGE in der Lehre zu befragen. Die Auswahl dieser Methode erfolgte auf Basis der bestehenden Literatur, nach welcher Fokusgruppen als ressourcenschonendes Erhebungsinstrument gut zur Meinungs- und Akzeptanzanalyse geeignet sind (Schulz et al. 2012). An jedem der Fokusgruppengespräche ließen wir sechs Studierende teilnehmen, da die empfohlene Gruppengröße bei etwa sechs bis zehn Teilnehmern liegt (Henseling et al. 2006). Im Gegensatz zu Einzelinterviews kommt es bei Gruppengesprächen zur Entwicklung von gruppendynamischen Effekten, welche sich positiv auf die Diskussionen der Teilnehmer auswirken und eine synergetische Qualität sichern (Henseling et al. 2006; Schulz et al. 2012). Die Tatsache, dass es sich im Falle dieser Arbeit um eine homogene Gruppe (Medizinstudierende des Repetitionsmoduls 6.x) handelte, die sich dennoch in mindestens einem Merkmal, dem Geschlecht, unterschied, erleichterte den Studierenden den Aufbau einer positiven Gruppendynamik (Henseling et al. 2006). Auch wenn laut Morgan (1996) drei bis fünf Fokusgruppen pro Projekt durchgeführt werden sollten, richtet sich die Anzahl der Fokusgruppengespräche vielmehr nach dem Erreichen einer Sättigung (Henseling et al. 2006). Dementsprechend wurden im Rahmen der vorliegenden Arbeit drei Fokusgruppengespräche durchgeführt, weil die Studierenden im dritten Gespräch keine neuen Aspekte mehr in die Diskussion einbrachten.

Zwischen der Moderatorin und den Gesprächsteilnehmern bestand keinerlei soziale Verbindung, was eine Antwortverzerrung aufgrund „sozialer Erwünschtheit“, d. h. aufgrund der Gabe vermeintlich erwünschter Antworten im Sinne von gegebenen sozialen Normen (Ellerkmann 2003), weitestgehend verhinderte. Die Moderatorin führte die Diskussionen neutral und hielt die von Bohnsack und Przyborski (2010) aufgestellten Anforderungen an einen Moderator von Gruppendiskussionen (u. a. Ansprechen der gesamten Gruppe; keine Abgabe von inhaltlichen Stellungnahmen) ein. Trotz dessen ist eine Einflussnahme der Moderatorin auf den Verlauf und die Ergebnisse der Fokusgruppengespräche nicht vollständig auszuschließen. Als Struktur durch das Gespräch diente ihr ein Leitfaden, welcher laut Schulz et al. (2012) die Vergleichbarkeit der Ergebnisse mehrerer Fokusgruppengespräche bezüglich spezifischer Fragestellungen steigert.

Im Rahmen dieser Studie haben sich mit einem Stichprobenumfang von $n=18$ (davon 1 Teilnehmer aus der POLEMA-Studienkohorte und 17 Teilnehmer aus der EMERGE-Studienkohorte) insgesamt 11,8% aller Studierenden des Repetitionsmoduls 6.x bzw. 20,5% aller Probanden der EMERGE-Gruppe zur Teilnahme an den Fokusgruppengesprächen bereit erklärt. Diese Zahlen sind vergleichbar mit denen anderer qualitativer Erhebungen. In einer Studie von Kaczmarczyk et al. (2015) nahmen beispielsweise 11% derjenigen Studierenden ($n=281$), die im Rahmen einer Studie ein ‚*Serious Game*‘ in der Lehre des 3. oder 4. Medizinstudienjahres genutzt hatten, an Fokusgruppengesprächen teil. In einer anderen Studie von Bryce et al. (1998) umfasste das Studienkollektiv insgesamt 110 Medizinstudierende im 4. Studienjahr, von denen jeweils eine Hälfte in der Interventions- bzw. in der Kontrollgruppe war. An den durchgeführten Fokusgruppengesprächen nahmen 24 Studierende (21,8%) teil. Allerdings machten die Autoren keine Angaben dazu, wie viele Teilnehmer der Interventionsgruppe bzw. der Kontrollgruppe entstammten.

Es sollte angemerkt werden, dass die Studierenden im Rahmen der hier vorliegenden Studie einen Büchergutschein als Aufwandsentschädigung für die Teilnahme an den Fokusgruppengesprächen erhielten. Eine solche monetäre Entschädigung stellt für Teilnehmer mit geringem Einkommen (z. B. Studierende) einen Teilnahmeanreiz dar (Schulz et al. 2012). Entsprechend könnte dieser Aspekt die Repräsentativität der Stichprobe und damit die Generalisierbarkeit der Ergebnisse einschränken. Allerdings wird im Rahmen von Fokusgruppengesprächen generell nicht die Generierung repräsentativer oder allgemeingültiger Ergebnisse angestrebt, sondern die Identifizierung von Trends und Mustern bezüglich spezifischer Themengebiete (Henseling et al. 2006).

4.7 Studiendesign

Bei der hier vorliegenden Studie handelt es sich um eine prospektive, monozentrische, nicht randomisierte Interventionsstudie, die im Sommersemester 2016 an der UMG stattfand.

Die effektive Rücklaufquote betrug 73,2% (112 von 153 möglichen Studierenden) und war vergleichbar mit den Rücklaufquoten anderer Studien zum Einsatz von ‚*Serious Games*‘ in der medizinischen Lehre (z. B. 70,4% in einer Studie von Kononowicz et al. (2012) bzw. 58,5% in einer Studie von LeFlore et al. (2012)). In einigen publizierten Studien wurde allerdings lediglich der endgültige Stichprobenumfang angegeben ohne zu erläutern, wie viele Probanden potentiell an der Studie hätten teilnehmen dürfen oder wie viele Probanden aufgrund bestimmter Kriterien ausgeschlossen werden mussten (Carr et al. 1999; Wahlgren et al. 2006; Dankbaar et al. 2016a).

Limitierend handelt es sich um eine monozentrische Einzelstudie, an welcher lediglich Studierende des 6. klinischen Semesters teilnehmen konnten. Außerdem präsentierten sowohl alle EMERGE- als auch alle ‚Key Feature‘-Prüfungsfälle relevante Krankheitsbilder aus den Fachbereichen „Innere Medizin“ und „Notfallmedizin“. Aus diesen Gründen können die Ergebnisse dieser Studie nicht auf das gesamte Gebiet der Humanmedizin verallgemeinert werden. In Folgestudien sollte untersucht werden, ob sich die Ergebnisse auch auf Studierende anderer klinischer Semester und auf andere medizinische Fachbereiche übertragen lassen. Um eine größere Repräsentativität der Ergebnisse zu erzielen, wird zudem die Durchführung einer multizentrischen Studie vorgeschlagen. Vermutlich ist diese jedoch mit einem großen logistischen und technischen Aufwand verbunden, da aktuell nicht alle medizinischen Hochschulen mit 3D-fähigen Computerclients zum Spielen von EMERGE ausgestattet sind und gegebenenfalls zunächst in die technische Aufrüstung der beteiligten Hochschulen investiert werden müsste.

Ferner waren die Studiengruppen nicht randomisiert, da die Studierenden selbst über ihre Teilnahme an EMERGE oder POLEMA im Rahmen des Repetitionsmoduls 6.x entscheiden durften. Eine randomisierte Studie mit einer Kontrollgruppe wäre wünschenswert gewesen, um die Validität der Studie zu erhöhen. Es war allerdings ethisch nicht vertretbar, Studierende gegen ihren Willen zur Teilnahme an einem der beiden Lehrformate zu zwingen. Dies galt insbesondere für EMERGE, einem Lehrformat, das im Sommersemester 2016 erstmalig in die curriculare Lehre integriert wurde und dessen Effektivität bis zu diesem Zeitpunkt weitgehend unbekannt war. Möglicherweise war die freiwillige Teilnahme der Studierenden an EMERGE ein Ausdruck ihrer Präferenz bezüglich digitaler Ressourcen. Diese Vermutung wurde dadurch unterstützt, dass die EMERGE-Teilnehmer die Fragebogensaussagen „Beim Arbeiten mit Computern habe ich Spaß“ ($2,4 \pm 1,1$ vs. $3,1 \pm 1,6$; $p=0,012$) und „Ich finde, dass Computer beim Lernen hilfreich sind“ ($1,9 \pm 0,9$ vs. $2,4 \pm 1,2$; $p=0,012$) signifikant positiver bewerteten als die POLEMA-Teilnehmer. Es bleibt somit zu überprüfen, ob sich die in dieser Studie beobachteten Ergebnisse ebenso in einer randomisierten Studie nachweisen lassen, in welcher persönliche Präferenzen keine Rollen spielen.

Wie oben ausgeführt (Kapitel 4.2 und 4.5) wird für weitere Studien zudem empfohlen, einen Prätest einzuführen, um das vorherige Leistungsniveau der Studierenden evaluieren und Lernerfolge, die tatsächlich während bzw. durch die Intervention zustande gekommen sind, unterscheiden zu können.

4.8 Ausblick

Im folgenden Kapitel werden mögliche Anknüpfungspunkte für die weitere Forschung aufgezeigt. Unter anderem warf die vorliegende Studie die Frage auf, wie genau ‚*Serious Games*‘ am effektivsten in der Lehre eingesetzt werden können. Darüber hinaus sollten in weiteren Forschungsarbeiten Fragen bezüglich der Langzeitretention erworbener Kompetenzen und des Transfers von Wissen und Fähigkeiten aus der Virtualität in die Realität beantwortet werden.

4.8.1 Effektiver Einsatz von ‚*Serious Games*‘ in der Lehre

Aus der vorliegenden Forschungsarbeit ergab sich die Frage nach den Bedingungen, die den Einsatz von ‚*Serious Games*‘ in der Lehre wirksam machen können. In den Fokusgruppengesprächen diskutierten die Teilnehmer darüber, wie der Einsatz von EMERGE in der Lehre organisiert werden sollte (71 Codings), um eine möglichst große Effektivität hinsichtlich der Förderung von ‚*clinical reasoning*‘-Fähigkeiten zu erreichen. Einige Studierende schlugen vor, EMERGE in Kleingruppen gemeinsam zu spielen (23 Codings) und eine Nachbesprechung bestimmter Fälle in der Gruppe anzubieten (7 Codings). Zudem gaben die Fokusgruppengespräche Hinweise darauf, dass das Betreuungsverhältnis zwischen dem ärztlichen Betreuer und den Studierenden, das in dieser Studie etwa 1:23 betrug, ausgebaut werden sollte (9 Codings). Die Diskussion von Fragen mit dem ärztlichen Betreuer wurde von den Studierenden als sehr hilfreich empfunden (11 Codings), sie merkten aber an, dass der Betreuer aus Kapazitätsgründen nicht allen Studierenden im gewünschten Ausmaß zur Verfügung stehen konnte. Nicht zuletzt kritisierten die Studierenden in Übereinstimmung mit Probanden weiterer Studien, u. a. von Dankbaar et al. (2016a), das Feedback, welches sie nach Fallabschluss im Archivmenü einsehen konnten (47 Codings). Tatsächlich zeigten Studien, dass ein intensives Feedback für einen effektiven Lernprozess notwendig ist (Ende 1983; Burgess und Mellis 2015) und speziell im Hinblick auf ‚*Serious Games*‘ ein persönliches und ausführliches Feedback mit stärkeren Lernerfolgen assoziiert sein könnte als ein oberflächliches Feedback (Lyon et al. 1991). Zudem sind im von Garris et al. (2002) postulierten Spielzyklus (siehe Kapitel 1.4.6.2) die Schritte „Feedback“ und „Reflexion und Analyse der Spielereignisse“ (das sog. ‚*Debriefing*‘) enorm wichtig, um Spielerfahrungen in Lernerfahrungen umwandeln und mithilfe von ‚*Serious Games*‘ Lernziele erreichen zu können. In Folgestudien sollte untersucht werden, wie genau das Feedback und das ‚*Debriefing*‘ gestaltet werden müssen, um ein effektives Lernen mit ‚*Serious Games*‘ zu gewährleisten.

4.8.2 Langzeitretention erworbener Kompetenzen

In der vorliegenden Studie wurde die Abschlussprüfung wenige Tage nach dem letzten EMERGE- bzw. POLEMA-Seminar, d. h. noch innerhalb des sechswöchigen Repetitionsmoduls 6.x, durchgeführt. Auf Basis der Ergebnisse dieser Studie kann dementsprechend nur die kurzfristige Effektivität von EMERGE und POLEMA im Hinblick auf den Erwerb von ‚*clinical reasoning*‘-Fähigkeiten beurteilt werden.

Auch in publizierten Studien erfolgten die Posttests zur Evaluation der Effektivität von ‚*Serious Games*‘ häufig unmittelbar im Anschluss an die Interventionsphase (Youngblood et al. 2008; LeFlore et al. 2012; Dankbaar et al. 2016a) oder beispielsweise im Abstand von einer Woche (Carr et al. 1999). In diesem Zusammenhang stellt sich die Frage, ob die mithilfe von ‚*Serious Games*‘ erworbenen ‚*clinical reasoning*‘-Fähigkeiten ins Langzeitgedächtnis übertragen und dauerhaft abgerufen werden können. In einer Metaanalyse von Sitzmann (2011), auf die bereits in Kapitel 4.2.1 eingegangen wurde, zeigten Studierende, die mithilfe von ‚*Serious Games*‘ gelernt haben, eine um insgesamt 9% bessere Retention von Wissen und Fähigkeiten als Studierende, die traditionelle Lehre erhalten haben. Allerdings wurde auch in dieser Metaanalyse angemerkt, dass die Wissensretention in den meisten Studien bereits ein bis vier Wochen nach Ende der Trainingsphase überprüft wurde. Ferner weisen entsprechende Studien bedeutende Limitationen auf, beispielsweise eine kleine Stichprobengröße von nur acht Probanden (Vukanovic-Criley et al. 2008).

In einer Studie von Diehl et al. (2015), siehe Kapitel 4.2.1, zeigten Studierende, die entweder mithilfe eines ‚*Serious Game*‘ oder mithilfe von traditioneller Lehre gelernt haben, drei Monate nach Abschluss der Interventionsphase eine gleich gute Retention von Wissen und Fähigkeiten zur Insulintherapie. Zwischen einer unmittelbar nach der Trainingsphase durchgeführten Prüfung und einer drei Monate später stattfindenden Erhebung verringerte sich die Leistung der Interventionsgruppe zwar von 90% auf 78% ($p < 0,001$) und die der Kontrollgruppe von 89% auf 80% ($p < 0,001$). Beide Gruppen zeigten aber auch noch drei Monate nach Abschluss der Trainingsphase signifikant bessere Leistungen als vor Beginn der Interventionsphase ($p < 0,001$). Für den Einsatz virtueller Patienten gaben Studien sogar Hinweise darauf, dass sich erworbene Kompetenzen, beispielsweise bezüglich der kardialen Untersuchung, nach Abschluss der Interventionsphase über mehrere Monate weiter verbesserten (Vukanovic-Criley et al. 2008).

Es wird vorgeschlagen, in weiteren Forschungsarbeiten die Langzeitretention von Kompetenzen, die mithilfe eines ‚*Serious Game*‘ trainiert wurden, durch Postmessungen zu

überprüfen. Diese sollten im Abstand von mehreren Monaten bzw. Jahren zur Interventionsphase durchgeführt werden und eine repräsentative Stichprobe einschließen.

4.8.3 Transfer von Wissen und Fähigkeiten aus der Virtualität in die Realität

Schließlich haben bislang nur wenige Studien untersucht, ob durch ‚*Serious Games*‘ virtuell trainierte ‚*clinical reasoning*‘-Fähigkeiten tatsächlich in die klinische Praxis transferiert werden können (Girard et al. 2013). Insofern ist unklar, ob ein entsprechendes Training tatsächlich nützlich für die klinische Praxis ist und die Kompetenz verbessert, bei der Behandlung realer Patienten Diagnose- und Therapieentscheidungen zu fällen und effektive Behandlungsabläufe zu wählen (Vukanovic-Criley et al. 2008).

Für die Behandlung von Simulations- bzw. Schauspielpatienten konnten entsprechende Lernerfolge durchaus bereits bestätigt werden. In einer Studie von LeFlore et al. (2012) führten beispielsweise Krankenpflegestudierende, die zuvor mithilfe eines ‚*Serious Game*‘ gelernt hatten, die zur Behandlung respiratorischer Symptome zweier Schauspielpatienten notwendigen Schritte signifikant schneller durch als Studierende, die das ‚*Serious Game*‘ zuvor nicht genutzt hatten ($p=0,001$). Konkret handelte es sich allerdings um einen Zeitunterschied von etwa einer Minute, welcher in der klinischen Realität sehr gering ist und lediglich in wenigen, kritischen Krankheitsfällen einen wesentlichen Unterschied im Hinblick auf das Behandlungsergebnis hervorruft. Weiterhin wandten Studierende, die im Rahmen einer Studie von Dankbaar et al. (2016a) das ‚*Serious Game*‘ ‚abcdeSIM‘ genutzt hatten, ihre praktischen, klinischen und kommunikativen Kompetenzen bei der Behandlung von Simulationspuppen gleich gut an wie Studierende, die zuvor ein Online-Lernmodul ohne Spielcharakter genutzt hatten. Auch wenn in diesen Studien versucht wurde, ein möglichst realistisches Szenario zu konstruieren, wurde der Praxistransfer erworbener Kompetenzen lediglich in einer simulierten Prüfungssituation untersucht. Die Frage, ob durch ‚*Serious Games*‘ in der Virtualität trainierte ‚*clinical reasoning*‘-Fähigkeiten tatsächlich erfolgreich in die praktische Realität transferiert und während der klinischen Tätigkeit erfolgreich angewendet werden können, sollte in zukünftigen Forschungsarbeiten beantwortet werden.

5 Zusammenfassung

Für eine effektive Patientenversorgung müssen Ärzte über mindestens zwei Schlüsselkompetenzen verfügen. Erstens müssen sie in der Lage sein, auf Grundlage einer Anamnese eine Verdachtsdiagnose zu formulieren, die zu ihrer Bestätigung erforderlichen Untersuchungen zu veranlassen und eine adäquate initiale Therapie einzuleiten. Die zugrundeliegenden Denkprozesse werden als ‚*clinical reasoning*‘ bezeichnet. Zweitens müssen sie mit Stress umgehen und bei der Behandlung mehrerer Patienten Prioritäten setzen können. Die Vermittlung bzw. der Erwerb dieser Kompetenzen gehört zu den wichtigsten und zugleich schwierigsten Aufgaben in der medizinischen Ausbildung.

Bislang standen zur Vermittlung derartiger Lernziele vor allem ressourcenintensive Lehrformate, u. a. Kleingruppenunterricht und Unterricht am Krankenbett, zur Verfügung. Allerdings konnte diese Lehre weder standardisiert werden, noch bot sie ausreichende Möglichkeit zur Diskussion seltener Erkrankungen und dringlicher Notfälle. Die zunehmende Verfügbarkeit digitaler Infrastruktur an medizinischen Fakultäten ermöglicht nun den Einsatz innovativer Lehrformate, wie zum Beispiel ‚*Serious Games*‘. Hierbei handelt es sich um didaktisch fundierte Lernspiele zum Training komplexer kognitiver Funktionen, in denen Studierende beispielsweise mit virtuellen Patienten konfrontiert werden, um ‚*clinical reasoning*‘ zu fördern. Obwohl Vorarbeiten zeigen, dass Studierende hochgradig motiviert sind, ‚*Serious Games*‘ zu nutzen und diese grundsätzlich zur Vermittlung ärztlicher Kompetenzen geeignet sind, existieren bislang kaum wissenschaftlich fundierte Daten zur Effektivität von ‚*Serious Games*‘ im Vergleich zu traditionellen Lehrformaten.

Im Rahmen dieser Forschungsarbeit wurde im Sommersemester 2016 die virtuelle Notaufnahme EMERGE, in welcher bis zu 10 virtuelle Patienten gleichzeitig behandelt werden müssen, in die curriculare Lehre der UMG integriert und ihre Effektivität im Hinblick auf das Training von ‚*clinical reasoning*‘ mit der Effektivität des problemorientierten Kleingruppenunterrichts POLEMA verglichen.

In POLEMA (n=34) wurden in kleinen Gruppen fünf verschiedenen Patientenfälle detailliert und nacheinander erarbeitet, wohingegen die Studierenden in EMERGE (n=78) bis zu zehn verschiedene Fälle pro Termin am eigenen Computer bearbeiteten. Innerhalb der EMERGE-Gruppe gab es drei Untergruppen EMERGE A, B und AB, die unterschiedlich stark gegenüber verschiedenen Fällen aus vier Fachbereichen der Inneren Medizin exponiert wurden. Anhand einer formativen Abschlussprüfung, bestehend aus sechs ‚*Key Feature*‘-Fällen und vier EMERGE-Fällen, wurden die ‚*clinical reasoning*‘-Fähigkeiten der Studierenden

im Anschluss an die Trainingsphase evaluiert. Zudem wurden im Rahmen der Datenerhebung Fragebögen und Fokusgruppengespräche eingesetzt.

Die Ergebnisse der ‚*Key Feature*‘-Prüfung zeigen, dass die Effektivität von EMERGE und POLEMA im Hinblick auf das Training differentialdiagnostischer und –therapeutischer Kompetenzen als mindestens gleichwertig einzustufen ist. So erreichten die EMERGE-Teilnehmer eine signifikant höhere Gesamtpunktzahl als die POLEMA-Teilnehmer ($61,8 \pm 14,5\%$ vs. $54,4 \pm 14,4\%$; $p=0,015$). Bezüglich Inhalten, die zuvor in beiden Lehrformaten thematisiert worden waren, ergab sich kein signifikanter Leistungsunterschied ($61,0 \pm 18,9\%$ vs. $61,5 \pm 21,1\%$; $p=0,831$). Der Vorteil von EMERGE liegt hierbei darin, dass den Studierenden innerhalb derselben Lernzeit mindestens vier Mal so viele Fälle präsentiert worden waren wie den POLEMA-Teilnehmern.

In der EMERGE-Abschlussprüfung erzielten die EMERGE-Teilnehmer in allen vier Fällen eine signifikant höhere Gesamtpunktzahl als die POLEMA-Teilnehmer. Ein signifikanter Vorteil der Computersimulation gegenüber dem problemorientierten Kleingruppenunterricht zeigte sich ferner in vier von sieben Kategorien bzgl. inhaltlicher und prozeduraler Fähigkeiten. Im Gegensatz dazu schnitten die drei EMERGE-Gruppen unabhängig vom Ausmaß ihrer Exposition gegenüber Fällen bestimmter Fachbereiche während der Trainingsphase in allen vier Fällen gleich gut ab. Dies legt nahe, dass die in EMERGE erlernten ‚*clinical reasoning*‘-Fähigkeiten nicht fallspezifisch sind, sondern erfolgreich auf unbekannte Fälle übertragen werden können.

Es ergaben sich zudem Hinweise darauf, dass das Training von ‚*clinical reasoning*‘ mithilfe einer virtuellen Notaufnahme das Stressempfinden bei der simultanen Behandlung mehrerer Patienten vermindert, denn die EMERGE-Teilnehmer gaben in den Fragebögen an, sich bei der gleichzeitigen Behandlung virtueller Patienten signifikant weniger gestresst zu fühlen als die POLEMA-Teilnehmer ($3,3 \pm 1,6$ vs. $4,3 \pm 1,4$; $p=0,001$).

Insgesamt ist EMERGE aus Sicht der Studierenden gut zur Vermittlung ärztlicher Kompetenzen und zum Training differentialdiagnostischer und –therapeutischer Kompetenzen geeignet. Zukünftige Studien sollten unter anderem untersuchen, wie genau ‚*Serious Games*‘ am effektivsten im Medizinstudium eingesetzt werden können und ob ein Transfer erlernter Kompetenzen aus der Virtualität in die Realität gelingt.

6 Anhang

Anhang 1: EMERGE Lerntagebuch aus dem Modulhandbuch des Repetitionsmoduls M6.x

EMERGE-Lerntagebuch

Name	_____
Datum	_____
Kurstag	_____

1) Welche Symptomkomplexe habe ich heute behandelt?

- | | |
|---------|---------|
| 1 _____ | 4 _____ |
| 2 _____ | 5 _____ |
| 3 _____ | 6 _____ |

2) Welche Materialien und Leitlinien habe ich heute benutzt?

3) Welche Aspekte waren für mich hierbei heute besonders wichtig?

4) Welche weiterführenden Fragen haben sich für mich heute ergeben?

5) Welche Fragen möchte ich bis zum nächsten EMERGE-Termin klären?

Anhang 2: Fragebogen (Beispiel)

EvaSys	Evaluation - EMERGE - Sommersemester 2016	
Universitätsmedizin Göttingen	Prof. Dr. med. Tobias Raupach, MME	:UMG
Bereich Medizindidaktik & Ausbildungsforschung	Evaluation EMERGE SoSe 2016	

Bitte so markieren: Bitte verwenden Sie einen Kugelschreiber oder nicht zu starken Filzstift. Dieser Fragebogen wird maschinell erfasst.
Korrektur: Bitte beachten Sie im Interesse einer optimalen Datenerfassung die links gegebenen Hinweise beim Ausfüllen.

Allgemeine Angaben und Vorerfahrungen

Sehr geehrte Studierende,
Im Rahmen des neuen Moduls "Klug entscheiden in der Medizin" wurde erstmals die virtuelle Notaufnahme EMERGE in der klinischen Lehre eingesetzt. Ihr Feedback ist für den zukünftigen Einsatz von EMERGE in der klinischen Lehre an der UMG von großer Bedeutung. Wir möchten Sie daher bitten, an dieser Evaluation teilzunehmen.

Bitte geben Sie Ihr Geschlecht an weiblich männlich
Bitte geben Sie Ihr Alter an

Bitte geben Sie an, an welcher Lehrform Sie im Modul 6.x teilgenommen haben POLEMA EMERGE
Haben Sie schon einmal in einer Notaufnahme gearbeitet (egal in welcher Funktion)? ja nein
Haben Sie schon einmal im Rettungsdienst gearbeitet (egal in welcher Funktion)? ja nein

Nutzung digitaler Medien

Wie schätzen Sie Ihre Kenntnisse im Umgang mit Computern ein? überhaupt keine Anfängerin/Anfänger Fortgeschrittene/
 Expertin/Experte
Wie viele Stunden verbringen Sie pro Tag insgesamt am Computer? <1 h pro Tag 1-2 h pro Tag 2-3 h pro Tag
 3-4 h pro Tag 4-5 h pro Tag > 5 h pro Tag
Wie regelmäßig spielen Sie Computerspiele? nie etwa einmal pro Jahr etwa einmal pro Monat
 etwa einmal pro Woche jeden Tag
Haben Sie schon einmal ein anderes Lernspiel („serious game“) als EMERGE genutzt? ja nein

Bitte geben Sie bei den nachfolgenden drei Fragen auf der 6-stufigen Skala an, inwiefern die jeweilige Aussage zutrifft.

„Beim Arbeiten mit Computern habe ich Spaß.“	trifft voll zu	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu
„Ich finde, dass Computer beim Lernen hilfreich sind.“	trifft voll zu	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu
„Mir fällt es meist leicht, zu erlernen, wie man ein neues Computerprogramm verwendet.“	trifft voll zu	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu
„Beim Spielen der heutigen EMERGE-Abschlusssitzung habe ich mich gestresst geföhlt.“	trifft voll zu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu

Motivation (nur für die EMERGE-Gruppe)

Bitte geben Sie bei den nachfolgenden acht Fragen auf der 6-stufigen Skala an, inwiefern die jeweilige Aussage zutrifft.

„Ich habe mich zur Teilnahme an EMERGE entschieden ...“

„... weil ich mich für den Einsatz von Lernspielen im Medizinstudium interessiere.“	trifft voll zu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu
„... weil ich lieber selbstständig Fälle bearbeite als diese in einer Kleingruppe zu besprechen.“	trifft voll zu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	trifft überhaupt nicht zu

F12955U460440448P1PL0V0 17.05.2016, Seite 1/2

EvaSys	Evaluation - EMERGE - Sommersemester 2016	Electric Paper	
Motivation (nur für die EMERGE-Gruppe) [Fortsetzung]			
„... weil ich dadurch einen geringeren Arbeitsaufwand als bei POLEMA erwartet habe.“	trifft voll zu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>	trifft überhaupt nicht zu
„... weil ich der EMERGE-Gruppe zugelost wurde.“	trifft voll zu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>	trifft überhaupt nicht zu
„... weil ich gerne am Computer arbeite.“	trifft voll zu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„... um differentialdiagnostische Fertigkeiten zu trainieren.“	trifft voll zu	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„... um mein therapeutisches Wissen zu erweitern.“	trifft voll zu	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„... um zu lernen, die Dringlichkeit klinischer Situationen richtig einzuschätzen.“	trifft voll zu	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
Inhaltliche Bewertung der virtuellen Notaufnahme (nur für die EMERGE-Gruppe)			
Bitte geben Sie bei den nachfolgenden sechs Fragen auf der 6-stufigen Skala an, inwiefern die jeweilige Aussage zutrifft.			
„Ich fühlte mich bei der gleichzeitigen Behandlung mehrerer Patienten während den EMERGE-Übungssitzungen gestresst.“	trifft voll zu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„Ich halte EMERGE zur Vermittlung ärztlicher Kompetenzen im Medizinstudium geeignet.“	trifft voll zu	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„Durch das Spielen von EMERGE habe ich meine Kompetenz gesteigert, die zu einem Symptomkomplex gehörenden Differentialdiagnosen zu kennen.“	trifft voll zu	<input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„Durch das Spielen von EMERGE habe ich gelernt, die Dringlichkeit einer klinischen Situation abzuschätzen.“	trifft voll zu	<input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„Durch das Spielen von EMERGE habe ich gelernt, die Wertigkeit verschiedener diagnostischer Schritte einschätzen zu können.“	trifft voll zu	<input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu
„Durch das Spielen von EMERGE habe ich gelernt, die bei einer Diagnose relevanten therapeutischen Schritte einzuleiten.“	trifft voll zu	<input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	trifft überhaupt nicht zu

Vielen Dank für Ihre Teilnahme!

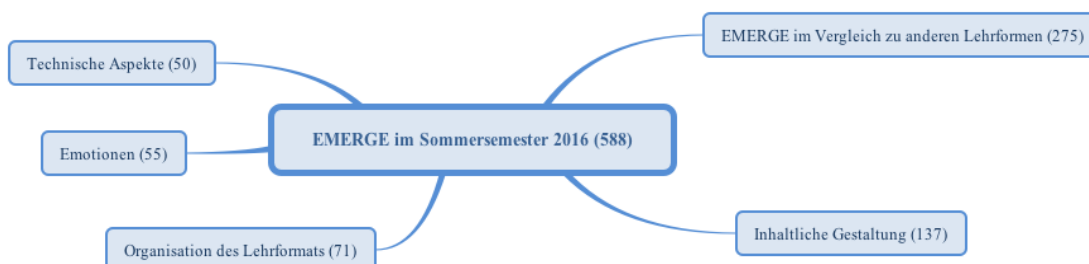


Anhang 3: Fragenkatalog für die Fokusgruppengespräche

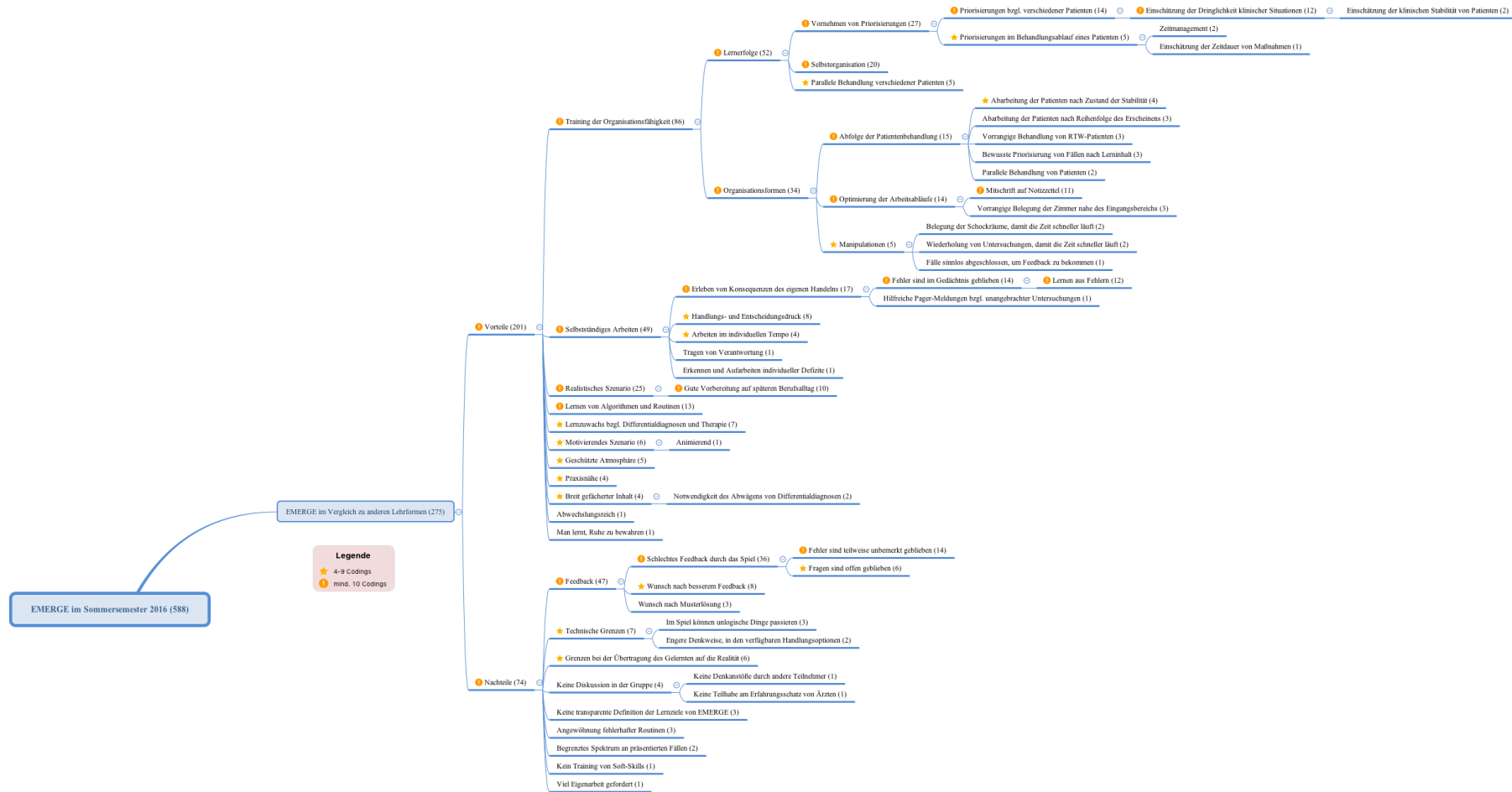
Leitfragen für das Fokusgruppengespräch zu EMERGE im Sommersemester 2016

1. **Was ist Euch von den EMERGE-Seminaren besonders in Erinnerung geblieben?**
 - a. Was war besonders hervorstechend? Warum?
 - b. Hat sich Eure Sichtweise auf EMERGE im Lauf der Zeit verändert?
 - c. Warum hat sich Eure Sichtweise verändert?
2. **Wie lief ein typischer EMERGE-Termin bei Euch ab?**
 - a. Was habt Ihr aus den EMERGE-Seminaren mitgenommen?
 - b. Wann hattet Ihr das Gefühl, während EMERGE besonders effektiv zu arbeiten?
 - c. Habt Ihr Hilfsmittel verwendet?
 - d. Wie habt Ihr Euch auf die EMERGE-Seminare vorbereitet?
 - a. Wie habt Ihr die EMERGE-Seminare nachbereitet?
3. **Welche Vor- und Nachteile seht Ihr zu anderen Lehrformen im Studium?**
 - a. **Gibt es Dinge, die Ihr bei EMERGE besser trainieren konntet als in anderen Lehrveranstaltungen?**
4. **Würdet Ihr etwas verändern?**
5. **Wie würdet Ihr ein EMERGE-Seminar gestalten?**
 - a. Was ist für die sinnvolle Gestaltung eines EMERGE-Seminars notwendig?
 - b. Was ist bei der zeitlichen Gestaltung eines EMERGE-Seminars zu beachten?
 - c. Wäre es lehrreich, anderen beim EMERGE-Spielen zuzusehen?
6. **Welche Emotionen habt Ihr während der Spielsitzungen durchlebt?**
 - a. Wodurch wurde die Emotion ausgelöst?
 - b. Wie seid Ihr mit den Emotionen umgegangen?
 - c. Wie bewertet Ihr die erfahrenen Emotionen? Waren sie hilfreich oder hinderlich?
7. **Was habt Ihr für Euren späteren Berufsalltag lernen können?**
 - a. Wie hat sich Eure Einschätzung zur Dringlichkeit klinischer Situationen verändert?
 - b. Wie realistisch findet Ihr das Spiel?
8. **Inwieweit hat EMERGE zu dem Repetitionscharakter des Moduls 6.x gepasst?**
 - a. Was habt Ihr über verschiedene Differentialdiagnosen verschiedener Symptomkomplexe dazugelernt?
 - b. Was habt Ihr über diagnosegeleitete Therapien dazugelernt?

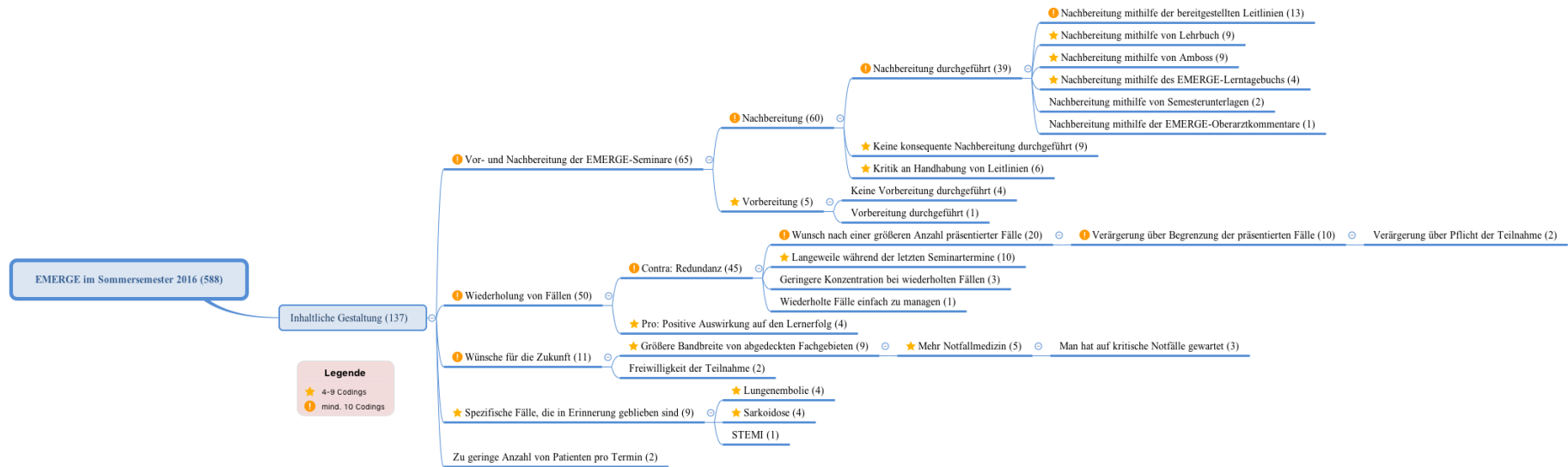
Anhang 4: MindMap der Auswertung der Fokusgruppengespräche



Anhang 5: MindMap „EMERGE im Vergleich zu anderen Lehrformen“



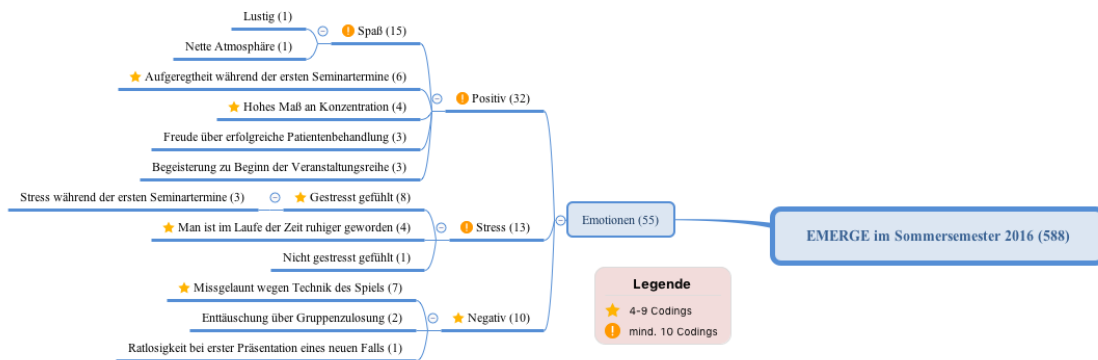
Anhang 6: MindMap „Inhaltliche Gestaltung“



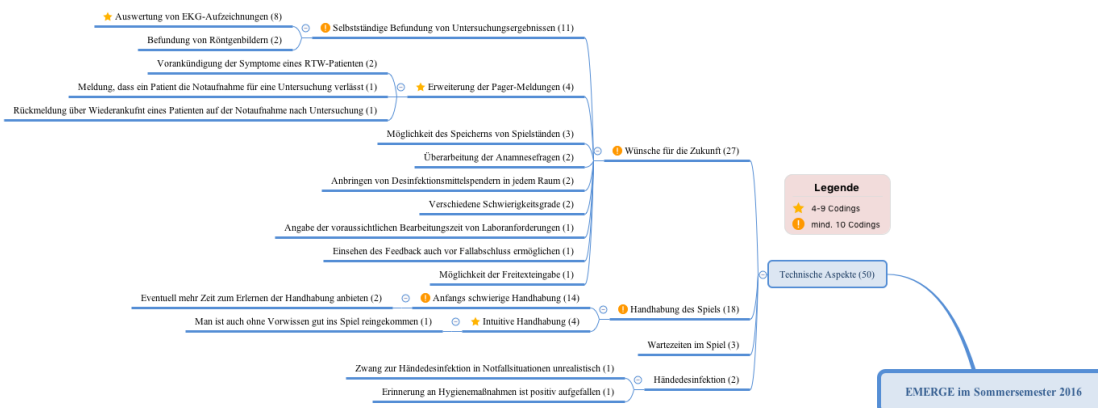
Anhang 7: MindMap „Organisation des Lehrformats“



Anhang 8: MindMap „Emotionen“



Anhang 9: MindMap „Technische Aspekte“



7 Literaturverzeichnis

Aldeen AZ, Gisondi MA (2006): Bedside Teaching in the Emergency Department. *Acad Emerg Med* 13, 860–866

Anziani H, Durham J, Moore U (2008): The relationship between formative and summative assessment of undergraduates in oral surgery: Formative and summative assessment of undergraduates. *Eur J Dent Educ* 12, 233–238

Ark TK, Brooks LR, Eva KW (2006): Giving learners the best of both worlds: Do clinical teachers need to guard against teaching pattern recognition to novices? *Acad Med* 81, 405–409

Barnes K, Marateo RC, Ferris SP (2007): Teaching and learning with the net generation. *Innov J Online Educ* 3, 1

Barrows HS (1993): An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med* 68, 443–453

Bearman M (2003): Is virtual the same as real? Medical students' experiences of a virtual patient. *Acad Med* 78, 538–545

Biggs J (1996): Enhancing teaching through constructive alignment. *High Educ* 32, 347–364

Bloice MD, Simonic K-M, Holzinger A (2014): Casebook: a virtual patient iPad application for teaching decision-making through the use of electronic health records. *BMC Med Inform Decis Mak* 14, 66

Bloom BS: *Taxonomy of Educational Objectives Book 1: Cognitive Domain*. 2nd edition; Addison Wesley Publishing Company, London 1984

Bohnsack R, Przyborski A: Diskursorganisation, Gesprächsanalyse und die Methode der Gruppendiskussion; in: *Das Gruppendiskussionsverfahren in der Forschungspraxis*, 2., vollständig überarbeitete Auflage; hrsg. v. Bohnsack R, Przyborski A, Schäffer B; Budrich, Opladen 2010, 233–248

Bowen JL (2006): Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med* 355, 2217–2225

Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, Collett T, de Bere SR (2010): The transition from medical student to junior doctor: today's experiences of tomorrow's doctors. *Med Educ* 44, 449–458

Bryce D, King N, Graebner C, Myers J (1998): Evaluation of a diagnostic reasoning program (DxR): Exploring student perceptions and addressing faculty concerns. *J Interact Media Educ* 98, 451–455

Burgess A, Mellis C (2015): Feedback and assessment for clinical placements: achieving the right balance. *Adv Med Educ Pract* 6, 373–381

- Carr MM, Reznick RK, Brown DH (1999): Comparison of computer-assisted instruction and seminar instruction to acquire psychomotor and cognitive knowledge of epistaxis management. *Otolaryngol Neck Surg* 121, 430–434
- Cave J, Goldacre M, Lambert T, Woolf K, Jones A, Dacre J (2007): Newly qualified doctors' views about whether their medical school had trained them well: questionnaire surveys. *BMC Med Educ* 7, 38
- Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP (2010): The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ* 15, 695–715
- Cohen S, Kamarck T, Mermelstein R (1983): A Global Measure of Perceived Stress. *J Health Soc Behav* 24, 385–396
- Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM (2012): A systematic literature review of empirical evidence on computer games and serious games. *Comput Educ* 59, 661–686
- Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM (2008): Internet-based learning in the health professions: a meta-analysis. *JAMA* 300, 1181–1196
- Cook DA, Erwin PJ, Triola MM (2010): Computerized Virtual Patients in Health Professions Education: A Systematic Review and Meta-Analysis. *Acad Med* 85, 1589–1602
- Cordova DI, Lepper MR (1996): Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *J Educ Psychol* 88, 715
- Croskerry P (2009): A Universal Model of Diagnostic Reasoning. *Acad Med* 84
- Dankbaar M, Alsmas J, Jansen EE, van Merriënboer JJ, van Saase JL, Schuit SC (2016a): An experimental study on the effects of a simulation game on students' clinical cognitive skills and motivation. *Adv Health Sci Educ* 21, 505–521
- Dankbaar ME, Roozeboom MB, Oprins EA, Rutten F, van Merriënboer JJ, van Saase JL, Schuit SC (2016b): Preparing Residents Effectively in Emergency Skills Training With a Serious Game: *Simul Healthc* 12, 9–16
- Dankbaar ME, Richters O, Kalkman CJ, Prins G, ten Cate OT, van Merriënboer JJ, Schuit SC (2017): Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness. *BMC Med Educ* 17, 30
- Davis MH, Harden RM (1999): AMEE Medical Education Guide No. 15: Problem-based learning: a practical guide. *Med Teach* 21, 130–140
- Des Marchais JE (1993): A student-centred, problem-based curriculum: 5 years' experience. *CMAJ Can Med Assoc J* 148, 1567
- Diehl LA, Gordan PA, Esteves RZ, Coelho ICMM (2015): Effectiveness of a serious game for medical education on insulin therapy: a pilot study. *Arch Endocrinol Metab* 59, 470–473

-
- Ellerkmann K: Soziale Erwünschtheit in der empirischen Sozialforschung. GRIN Verlag, München 2003
- Ende J (1983): Feedback in clinical medical education. *JAMA* 250, 777–781
- Epstein RM, Hundert EM (2002): Defining and assessing professional competence. *Jama* 287, 226–235
- Er E, Öden M, Arifoglu, A. (2009): A blended e-learning environment: A model proposition for integration of asynchronous and synchronous e-learning. *Int J Learn* 16, 449–460
- Ericsson KA (2004): Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 79, 70–81
- Eva KW (2005): What every teacher needs to know about clinical reasoning. *Med Educ* 39, 98–106
- Evans J (2003): In two minds: dual-process accounts of reasoning. *Trends Cogn Sci* 7, 454–459
- Evans J, Stanovich KE (2013): Dual-Process Theories of Higher Cognition Advancing the Debate. *Perspect Psychol Sci* 8, 223–241
- Farmer EA, Page G (2005): A practical guide to assessing clinical decision-making skills using the key features approach. *Med Educ* 39, 1188–1194
- Finucane PM, Johnson SM, Prideaux DJ (1998): Problem-based learning: its rationale and efficacy. *Med J Aust* 168, 445–447
- Fischer MR, Kopp V, Holzer M, Ruderich F, Jünger J (2005): A modified electronic key feature examination for undergraduate medical students: validation threats and opportunities. *Med Teach* 27, 450–455
- Frankish K, Evans J: The duality of mind: an historical perspective; in: *In two minds: Dual processes and beyond*; hrsg. v. Frankisch K, Evans J; Oxford University Press, Oxford 2009, 1–29
- Frenzel AC, Götz T, Pekrun R: Emotionen; in: *Pädagogische Psychologie*; hrsg. v. Wild E, Möller; Springer Medizin-Verlag, Heidelberg 2009, 205–231
- Friedman CP (2000): The marvelous medical education machine or how medical education can be unstuck in time. *Med Teach* 22, 496–502
- Garris R, Ahlers R, Driskell JE (2002): Games, Motivation, and Learning: A Research and Practice Model. *Simul Gaming* 33, 441–467
- Garrison DR, Kanuka H (2004): Blended learning: Uncovering its transformative potential in higher education. *Internet High Educ* 7, 95–105
- Gierk B, Harendza S (2012): Patient selection for bedside teaching: inclusion and exclusion criteria used by teachers. *Med Educ* 46, 228–233

- Girard C, Ecalle J, Magnan A (2013): Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *J Comput Assist Learn* 29, 207–219
- Graafland M, Schraagen JM, Schijven MP (2012): Systematic review of serious games for medical education and surgical skills training. *Br J Surg* 99, 1322–1330
- Gross R, Löffler M: *Prinzipien der Medizin: Eine Übersicht ihrer Grundlagen und Methoden*. Springer, Berlin 1997
- Güntert A: *Approbationsordnung für Ärzte (ÄAppO), Bundesärztleordnung (BÄO): mit Erläuterungen und praktischen Hinweisen*. Deutscher Ärzteverlag, Köln 2003
- Harden RM (1999): AMEE Guide No. 14: Outcome-based education: Part 1-An introduction to outcome-based education. *Med Teach* 21, 7–14
- Harendza S, Alofs L, Huiskes J, Wijnen-Meijer M (2013): Ordering patterns for laboratory and radiology tests by students from different undergraduate medical curricula. *BMC Med Educ* 13, 109
- Henseling C, Hahn T, Nolting K: *Die Fokusgruppen-Methode als Instrument in der Umwelt- und Nachhaltigkeitsforschung*. IZT, Berlin 2006
- Hrynychak P, Glover Takahashi S, Nayer M (2014): Key-feature questions for assessment of clinical reasoning: a literature review. *Med Educ* 48, 870–883
- Hudson JN (2006): Formative assessment can be fun as well as educational. *AJP Adv Physiol Educ* 30, 33–37
- Issenberg SB, Mcgaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ (2005): Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 27, 10–28
- Jaques D (2003): ABC of learning and teaching in medicine. *BMJ* 326, 492–494
- Kaczmarczyk J, Davidson R, Bryden D, Haselden S, Vivekananda-Schmidt P (2015): Learning decision making through serious games. *Clin Teach* 13, 277–82
- Kahneman D (2002): Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel Prize Lect* 8, 351–401
- Kassirer JP (2010): Teaching clinical reasoning: case-based and coached. *Acad Med* 85, 1118–1124
- Kern D, Thomas P, Howard D, Bass E: *Curriculum Development for Medical Education. A Six- Step- Approach*. John Hopkins University Press, Baltimore 1998
- Kiili K (2005): Digital game-based learning: Towards an experiential gaming model. *Internet High Educ* 8, 13–24
- Knight JF, Carley S, Tregunna B, Jarvis S, Smithies R, de Freitas S, Dunwell I, Mackway-Jones K (2010): Serious gaming technology in major incident triage training: A pragmatic

controlled trial. *Resuscitation* 81, 1175–1179

Kononowicz AA, Krawczyk P, Cebula G, Dembkowska M, Drab E, Frączek B, Stachoń AJ, Andres J (2012): Effects of introducing a voluntary virtual patient module to a basic life support with an automated external defibrillator course: a randomised trial. *BMC Med Educ* 12, 41

Kopp V, Möltner A, Fischer MR (2006): Key-Feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. *GMS Z Med Ausbild* 23, Doc50

Krapp A (2005): Emotion und Lernen—Beiträge der Pädagogischen Psychologie. Einführung in den Thementeil. *ZfPäd* 51, 603–609

Kroenke K (1992): Attending rounds. *J Gen Intern Med* 7, 68–75

Kron F, Gjerde C, Sen A, Fetters M (2010): Medical student attitudes toward video games and related new media technologies in medical education. *BMC Med Educ* 10, 50

LaCombe MA (1997): On bedside teaching. *Ann Intern Med* 126, 217–220

Lane JL, Slavin S, Ziv A (2001): Simulation in Medical Education: A Review. *Simul Gaming* 32, 297–314

LeFlore JL, Anderson M, Zielke MA, Nelson KA, Thomas PE, Hardee G, John LD (2012): Can a Virtual Patient Trainer Teach Student Nurses How to Save Lives—Teaching Nursing Students About Pediatric Respiratory Diseases. *J Soc Simul Healthc* 7, 10–17

Ludwig S, Schuelper N, Brown J, Anders S, Raupach T (2018): How can we teach medical students to choose wisely? A randomised controlled cross-over study of video- versus text-based case scenarios. *BMC Med* 16, 107

Lyon HC, Healy JC, Bell JR, O'Donnell JF, Shultz EK, Wigton RS, Hirai F, Beck JR (1991): Significant efficiency findings while controlling for the frequent confounders of CAI research in the Planalyzer project's computer-based, self-paced, case-based programs in anemia and chest pain diagnosis. *J Med Syst* 15, 117–132

Mayring P: *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 12. Auflage; Beltz, Weinheim 2015

McLachlan JC (2006): The relationship between assessment and learning. *Med Educ* 40, 716–717

Miglioretti DL, Smith-Bindman R (2011): Overuse of computed tomography and associated risks. *Am Fam Physician* 83, 1252–1254

Miyakis S, Karamanof G, Lontos M, Mountokalakis TD (2006): Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgrad Med J* 82, 823–829

Morgan DL (1996): *Focus groups as qualitative research*. Sage Publ 16

-
- Nair BR, Coughlan JL, Hensley MJ (1998): Impediments to bed-side teaching. *Med Educ-Oxf* 32, 159–162
- Neufeld VR, Barrows HS (1974): The „McMaster Philosophy“: An Approach to Medical Education. *J Med Educ* 49, 1040–1050
- Nikendei C, Jünger J (2006): OSCE-praktische Tipps zur Implementierung einer klinisch-praktischen Prüfung. *GMS Z Med Ausbild* 23, 2006–23
- Norman G (2005): Research in clinical reasoning: past history and current trends. *Med Educ* 39, 418–427
- Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S (2016): The Causes of Errors in Clinical Reasoning: Cognitive Biases, Knowledge Deficits, and Dual Process Thinking. *Acad Med* 92, 23–30
- Page G, Bordage G (1995): The Medical Council of Canada’s Key Features Project: A More Valid Written Examination of Clinical Decision-making Skills. *Acad Med* 70, 104–110
- Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP, Howe NL, Ronckers CM, Rajaraman P, Craft AW (2012): Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *The Lancet* 380, 499–505
- Pekrun R (2014): Emotions and learning. *Educ Pract Ser* 24
- Peters M, ten Cate O (2014): Bedside teaching in medical education: a literature review. *Perspect Med Educ* 3, 76–88
- Prensky M (2001): Digital natives, digital immigrants part 1. *Horiz* 9, 1–6
- Qureshi Z, Maxwell S (2012): Has bedside teaching had its day? *Adv Health Sci Educ* 17, 301–304
- Raupach T, Brown J, Anders S, Hasenfuss G, Harendza S (2013): Summative assessments are more powerful drivers of student learning than resource intensive teaching formats. *BMC Med* 11, 61
- Raupach T, Andresen JC, Meyer K, Strobel L, Koziolk M, Jung W, Brown J, Anders S (2016): Test-enhanced learning of clinical reasoning: a crossover randomised trial. *Med Educ* 50, 711–720
- Reusser K (2005): Problemorientiertes Lernen–Tiefenstruktur, Gestaltungsformen, Wirkung. *Beitr Zur Lehrerinnen- Lehrerbildung* 23, 159–182
- Roediger HL, Karpicke JD (2006): The power of testing memory: Basic research and implications for educational practice. *Perspect Psychol Sci* 1, 181–210
- Roffi M, Patrono C, Collet J-P, Mueller C, Valgimigli M, Andreotti F, Bax JJ, Borger MA, Brotons C, Chew DP, et al. (2016): 2015 ESC Guidelines for the management of acute

coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J* 37, 267–315

Ross SM, Morrison GR, O'Dell JK (1989): Uses and effects of learner control of context and instructional support in computer-based instruction. *Educ Technol Res Dev* 37, 29–39

Rotthoff T, Baehring T, Dicken H-D, Fahron U, Richter B, Fischer MR, Scherbaum WA (2006): Comparison between Long-Menu and Open-Ended Questions in computerized medical assessments. A randomized controlled trial. *BMC Med Educ* 6, 50

Ruiz JG, Mintzer MJ, Leipzig RM (2006): The impact of e-learning in medical education. *Acad Med* 81, 207–212

Ryan RM, Deci EL (2000): Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemp Educ Psychol* 25, 54–67

Salam A, Siraj HH, Mohamad N, Das S, Rabeya Y (2011): Bedside teaching in undergraduate medical education: issues, strategies, and new models for better preparation of new generation doctors. *Iran J Med Sci* 36, 1

Schulz M, Mack B, Renn O: *Fokusgruppen in der empirischen Sozialwissenschaft*. VS Verlag für Sozialwissenschaften, Wiesbaden 2012

Schuwirth LWT, Verheggen MM, van der Vleuten CPM, Boshuizen HPA, Dinant GJ (2001): Do short cases elicit different thinking processes than factual knowledge questions do? *Med Educ* 35, 348–356

Shuell TJ (1986): Cognitive Conceptions of Learning. *Rev Educ Res* 56, 411

Shumway JM, Harden RM (2003): AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach* 25, 569–584

Sitzmann T (2011): A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Pers Psychol* 64, 489–528

Spady WG (1988): *Organizing for Results: The Basis of Authentic Restructuring and Reform*. *Educ Leadersh* 46, 4–8

Spencer JA, Jordan RK (1999): Learner centred approaches in medical education. 318, 1289–1283

Srinivasan M, Wilkes M, Stevenson F, Nguyen T, Slavin S (2007): Comparing problem-based learning with case-based learning: effects of a major curricular shift at two institutions. *Acad Med* 82, 74–82

Stanovich KE: *Who Is Rational?: Studies of individual Differences in Reasoning*. Psychology Press, Mahwah 1999

Stanovich KE: *The Robot's Rebellion: Finding Meaning in the Age of Darwin*. 1. Auflage; University Of Chicago Press, Chicago 2004

-
- Towle A, Cottrell D (1996): Self directed learning. *Arch Dis Child* 74, 357–359
- Van der Vleuten C, Newble DI (1995): How can we test clinical reasoning? *The Lancet* 345, 1032–1034
- Vash JH, Yunesian M, Shariati M, Keshvari A, Harirchi I (2007): Virtual patients in undergraduate surgery education: A randomized controlled study. *ANZ J Surg* 77, 54–59
- Vivekananda-Schmidt P, Psycholb C (2011): A Serious Game for Teaching Nursing Students Clinical Reasoning and Decision-Making Skills. *Nurs Inform* 2016 42, 905
- van der Vleuten CPM, Driessen EW (2014): What would happen to education if we take education evidence seriously? *Perspect Med Educ* 3, 222–232
- Vukanovic-Criley JM, Boker JR, Criley SR, Rajagopalan S, Criley JM (2008): Using Virtual Patients to Improve Cardiac Examination Competency in Medical Students. *Clin Cardiol* 31, 334–339
- Wahlgren C-F, Edelbring S, Fors U, Hindbeck H, Ståhle M (2006): Evaluation of an interactive case simulation system in dermatology and venereology for medical students. *BMC Med Educ* 6, 40
- Watts M, Becker WE (2008): A little more than chalk and talk: Results from a third national survey of teaching methods in undergraduate economics courses. *J Econ Educ* 39, 273–286
- Weinert FE: *Leistungsmessungen in Schulen*. 3. Auflage; Beltz, Weinheim 2014
- Winkens R, Dinant G-J (2002): Evidence base of clinical diagnosis: rational, cost effective use of investigations in clinical practice. *BMJ* 324, 783
- Wood DF (2003): Problem based learning. *Br Med J* 326, 328
- Young DW (1988): Improving laboratory usage: a review. *Postgrad Med J* 64, 283–289
- Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P (2008): Design, Development, and Evaluation of an Online Virtual Emergency Department for Training Trauma Teams: *J Soc Simul Healthc* 3, 146–153
- Ziv A, Wolpe PR, Small SD, Glick S (2003): Simulation-based medical education: an ethical imperative. *Acad Med* 78, 783–788

Internetquellen:

Association of American Medical Colleges. <https://aamc.org/publications>; Zugriff am 22.02.2017

Göttinger Lernzielkatalog. http://www.med.uni-goettingen.de/de/media/G1-2_lehre/lernzielkatalog.pdf; Zugriff am 29.10.2016

Studienordnung für den Studiengang Humanmedizin an der Georg-August-Universität Göttingen. http://www.med.uni-goettingen.de/de/media/G1-G2_lehre/studium_ordnungen_studienordnung_okt_2015.pdf; Zugriff am 28.02.2017

Universitätsmedizin Göttingen – Curriculum Klinik. <http://www.med.uni-goettingen.de/de/content/studium/16056.html>; Zugriff am 01.11.2016

Universitätsmedizin Göttingen – Klinischer Studienabschnitt Humanmedizin. <http://www.med.uni-goettingen.de/de/content/studium/939.html>; Zugriff am 29.10.2016

Danksagung

An dieser Stelle möchte ich mich herzlich bei meinem Doktorvater Prof. Dr. Tobias Raupach für die Überlassung des Dissertationsthemas, die engagierte Betreuung und Unterstützung während der gesamten Bearbeitungsphase und die wertvollen Erfahrungen, die ich während der Arbeit an meiner Dissertation und der Präsentation meiner Ergebnisse auf wissenschaftlichen Konferenzen sammeln durfte, bedanken.

Besonders bedanken möchte ich mich zudem bei Dr. Nikolai Schuelper für die Hilfe bei der Durchführung dieser Studie und die zahlreichen Diskussionen, Ratschläge und lieben Worte, die wesentlich zum Gelingen dieser Arbeit beitrugen.

Darüber hinaus gilt mein Dank Christian Münscher und allen weiteren Mitarbeitern der Lehr-IT des Bereichs Medizindidaktik und Ausbildungsforschung für die technische Unterstützung bei der Durchführung der EMERGE-Seminare und der elektronischen Abschlussprüfung.

Allen Mitarbeitern und Doktoranden des Bereichs Medizindidaktik und Ausbildungsforschung der UMG danke ich für ihre Hilfsbereitschaft und den Studierenden der UMG für ihre Teilnahme an dieser Studie.